Documenta Mathematica

Journal der
Deutschen Mathematiker-Vereinigung
Gegründet 1996

Band 21 · 2016

DOCUMENTA MATHEMATICA, Journal der Deutschen Mathematiker-Vereinigung, veröffentlicht Forschungsarbeiten aus allen mathematischen Gebieten und wird in traditioneller Weise referiert. Es wird indiziert durch Mathematical Reviews, Science Citation Index Expanded, Zentralblatt für Mathematik.

Artikel können als TEX-Dateien per E-Mail bei einem der Herausgeber eingereicht werden. Hinweise für die Vorbereitung der Artikel können unter der unten angegebenen WWW-Adresse gefunden werden.

DOCUMENTA MATHEMATICA, Journal der Deutschen Mathematiker-Vereinigung, publishes research manuscripts out of all mathematical fields and is refereed in the traditional manner. It is indexed in Mathematical Reviews, Science Citation Index Expanded, Zentralblatt für Mathematik.

Manuscripts should be submitted as TEX -files by e-mail to one of the editors. Hints for manuscript preparation can be found under the following web address.

`http://www.math.uni-bielefeld.de/documenta`

ISSN 1431-0635 (Print), ISSN 1431-0643 (Internet)

SPARC
LEADING EDGE

DOCUMENTA MATHEMATICA is a Leading Edge Partner of SPARC, the Scholarly Publishing and Academic Resource Coalition of the Association of Research Libraries (ARL), Washington DC, USA.

Address of Technical Managing Editor: Ulf Rehmann, Fakultät für Mathematik, Universität Bielefeld, Postfach 100131, D-33501 Bielefeld, Typesetting in TEX.

Documenta Mathematica

Band 21, 2016

On the Smoothness of Normalisers, the Subalgebra Structure of Modular Lie Algebras, and the Cohomology of Small Representations

Sebastian Herpel and David I. Stewart

Received: June 19, 2015
Revised: October 30, 2015

Communicated by Joachim Cuntz

Abstract. We provide results on the smoothness of normalisers in connected reductive algebraic groups G over fields k of positive characteristic p. Specifically we we give bounds on p which guarantee that normalisers of subalgebras of $\mathfrak{g}$ in G are smooth, i.e. so that the Lie algebras of these normalisers coincide with the infinitesimal normalisers.

One of our main tools is to exploit cohomology vanishing of small dimensional modules. Along the way, we obtain complete reducibility results for small dimensional modules in the spirit of similar results due to Jantzen, Guralnick, Serre and Bendel–Nakano–Pillen.

2010 Mathematics Subject Classification: Primary 17B45; Secondary 20G07, 20G10.
Keywords and Phrases: Smoothness, subgroup structure, complete reducibility, cohomology.

Contents

1 Introduction

Let G be an affine group scheme over an algebraically closed field k. We say G is smooth if $\dim \mathrm{Lie}(G) = \dim G$. A famous theorem of Cartier states that every affine group over a field of characteristic zero is smooth. Therefore, in this situation, the category of smooth group schemes is closed under the scheme-theoretic constructions of taking centres, centralisers, normalisers and transporters. However, Cartier's theorem fails rather comprehensively in positive characteristic. A classic example of a non-smooth algebraic group is the group scheme μ_p whose points are the pth roots of unity; this is not smooth over a field of characteristic p—its Lie algebra is 1-dimensional, but its k-points consist just of the identity element. Furthermore, since μ_p is also the scheme-theoretic centre of SL_p, the centre of this reductive[1] group is also not smooth over a field of characteristic p. This means that the group-theoretic centre of SL_p misses important infinitesimal information about the centre (for instance, the fact that SL_p is not adjoint).

Nonetheless, centralisers are usually smooth. For example, it is a critical result of Richardson [Ric67, Lem. 6.6], used extensively in the theory of nilpotent orbits, that the centraliser $G_e = C_G(e)$ of an element e of $\mathfrak{g} = \mathrm{Lie}(G)$ is smooth whenever p is a very good prime for G.[2] (Note that smoothness of the centraliser, or what is the same, the separability of the orbit map $G \to G \cdot e$

[1] We call a smooth algebraic group G reductive provided that $R_u(G^0) = 1$.

[2] Recall that p is good if the following holds: p is not 2 if G contains a factor not of type A, p is not 3 if G contains an exceptional factor and p is not 5 if G contains a factor of type E_8. The prime p is very good if it is good and it does not divide $n+1$ for any factor of G of type A_n.

can be restated as $\mathrm{Lie}(G_e(k)) = \mathfrak{c}_{\mathfrak{g}}(e)$.) In fact the centralisers of subgroup schemes of a connected reductive group G are usually smooth: work of Bate–Martin–Röhrle–Tange and the first author (cf. Proposition 3.1) gives precise information on the characteristic p of k, depending on the root datum of G, for centralisers of all subgroup schemes of G to be smooth. It suffices, for instance, for p to be very good for G. Furthermore, centralisers of all subgroup schemes of GL_n are smooth.

The situation for normalisers is much less straightforward, which may explain why results in this direction have been unforthcoming until now. For example, even when $G = \mathrm{GL}_n$, for any $n \geq 3$ and any $p > 0$ an arbitrary prime, there are connected smooth subgroups of G with non-smooth normalisers (see Lemma 11.11 below). In light of this situation, perhaps it is surprising that there are any general situations in which normalisers of subgroup schemes are smooth. However, we prove that for sufficiently large p depending on the connected reductive algebraic group G, (a) all normalisers of height one subgroup schemes (in fact the normalisers of all subspaces of the Lie algebra of G); and (b) all normalisers of connected reductive subgroups are indeed smooth. Theorem 3.2 makes (b) precise and the proof is a straightforward reduction to the case of centralisers. Our main result follows.

Theorem A. *There exists a constant $c = c(r)$ such that if $p > c$ and G is any connected reductive group of rank r then all normalisers $N_G(\mathfrak{h})$ of all subspaces $\mathfrak{h}$ of $\mathfrak{g}$ are smooth.*

More precisely, let d be the dimension of a minimal faithful representation of G. Then all normalisers of subspaces of $\mathfrak{g}$ are smooth provided that $p > 2^{2d}$. In particular, if $G = \mathrm{GL}_n$ we may take $p > 2^{2n}$.

Remarks 1.1. (a). Clearly, the constant $c(r)$ in the theorem may be defined as $2^{2d'}$ for d' the maximal dimension of a minimal faithful module of a connected reductive group of rank r.

(b). Note that the maximum is finite since there are only a finite number of isomorphism types of connected reductive groups of a given rank over an algebraically closed field k. Each of these arises by base change from a split reductive group defined over the integers, so one can consider the theorem as a statement that for a fixed group $G_{\mathbb{Z}}$, the conclusion holds for each reduction modulo p of $G_{\mathbb{Z}}$, whenever p is sufficiently large.

It is natural to ask if lower bounds for the constant c in Theorem A exist. In §11, we present a menagerie of examples where smoothness of normalisers fails; in particular, in Example 11.4 we give a p-subalgebra of $\mathfrak{gl}_{2n+12}$ with non-smooth normaliser whenever $p|F_n$, the nth Fibonacci number. Since $F_n \sim 1.6^n$ and infinitely many Fibonacci numbers are expected to be prime, we conclude that $c(G)$ should grow exponentially with the rank of G. In other words the bound on p in the theorem is likely to be of the right order.

The obstruction to finding linear bounds for c comes from the fact that one cannot, in general, lift the maximal tori of Lie-theoretic normalisers to group-theoretic normalisers. However, many interesting subalgebras of $\mathfrak{g}$ have normalisers which are generated by nilpotent elements (such as maximal semisimple subalgebras). Adding in this extra, natural hypothesis gives rise to much better bounds. In the following theorem let $h = h(G)$ denote the Coxeter number of (the root system Φ of) G. If Φ is reducible, then h is taken as the maximum over all components.

THEOREM B. *(i) Let G be a reductive algebraic group and let d be as in Theorem A. Suppose $p > d+1$. Then all normalisers $N_G(\mathfrak{h})$ of p-subalgebras $\mathfrak{h}$ are smooth whenever $\mathfrak{n}_{\mathfrak{g}}(\mathfrak{h})$ is generated by nilpotent elements. More precisely, the conclusion holds for normalisers generated by nilpotent elements when G is simple of classical type (that is, the root system of G is of A–D type) and $p > h+1$.*

(ii) Let $p > 2h-2$ for the connected reductive group G. Then the normalisers $N_G(\mathfrak{h})$ of all subspaces $\mathfrak{h}$ of $\mathfrak{g}$ are smooth whenever $\mathfrak{n}_{\mathfrak{g}}(\mathfrak{h})$ is generated by nilpotent elements.

Remarks 1.2. (a). The bounds in Theorem B(i) are tight when G is classical of type A, B or C: whenever $p \leq h+1$ the smallest irreducible representation of the first Witt algebra or its adjoint gives rise to a non-smooth normaliser which satisfies the hypotheses. Theorem B(i) is also tight for G_2, as it contains a copy of the Witt algebra as a maximal subalgebra when $p = 7$; more generally, the conclusion of Theorem B(i) fails for all exceptional algebraic groups when $p = h+1$ (see [HS16]).
(b). Suppose that k is not algebraically closed, and that G is a connected reductive algebraic group defined over k with a closed, k-defined subgroup-scheme H. Since smoothness is a geometric property, we have that $N_G(H)$ is smooth if and only if $N_{G_{\overline{k}}}(H_{\overline{k}})$ is smooth. Hence Theorems A and B give sufficient conditions for the smoothness of normalisers over general base fields.

In proving the theorems above we require several auxiliary results which may be of independent interest. The first is necessary in proving Theorem B(i).

THEOREM C. *Let $\mathfrak{g} = \mathrm{Lie}(G)$ for G a simply-connected classical algebraic group over an algebraically closed field k and let $p > 2$ be a very good prime for G. Then any maximal non-semisimple subalgebra of $\mathfrak{g}$ is parabolic.*

Remark 1.3. An announcement of a full classification of the maximal non-semisimple subalgebras of the Lie algebras of classical groups is given in [Ten87]. We provide a straightforward proof of the stated part in §7 below.

The proof of Theorem B(i) also uses a number of results on cohomology of low-dimensional modules. Such results have something of a history: in [Jan97] Jantzen proved that a module for a connected reductive algebraic group with $p \geq \dim V$ is completely reducible. Building on this, Guralnick tackled the case of finite simple groups in [Gur99]; this time one needs $p \geq \dim V + 2$ for the same conclusion. In a different direction, Serre proved in [Ser94] that if two semisimple modules V_1 and V_2 for an arbitrary group satisfy $\dim V_1 + \dim V_2 <$

$p+2$ then their tensor product is semisimple. Extending work of Bendel–Nakano–Pillen, we add analogues of these results for Lie algebras and Frobenius kernels of reductive algebraic groups tackling the 'crucial case' of a question of Serre [Ser94, Question 1.2] (though see Footnote 3 below). We summarise our results when G is simple into the following. The extensions to the case G is semisimple or reductive can be found in §8, where also can be found any unexplained terminology.

Theorem D. *Suppose G is a simple algebraic group and let G_r be its r-th Frobenius kernel with $\mathfrak{g}$ its Lie algebra. Let V be a k-vector space with* $\dim V \leq p$.

(a) Suppose V is a G_r-module. Then V is completely reducible unless $\dim V = p$*, and either G is of type A_1 or $p = 2$ and G is of type C_n. In the exceptional cases, V is known explicitly.*

(b) Suppose $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$ and V is a $\mathfrak{g}$-module. Then either V is completely reducible or $\dim V = p$*, G is of type A_1 and V is known explicitly.*

(c) Let $p > h$. Then $\mathrm{H}^2(\mathfrak{g}, L(\mu)) = 0$*, for all μ in the lowest alcove $C_{\mathbb{Z}}$, unless G is of type A_1 and $\mu = (p-2)$; or G is of type A_2 and $\mu = (p-3, 0)$ or $(0, p-3)$.*

(d) Suppose V and W are semisimple $\mathfrak{g}$-modules with $\dim V + \dim W < p+2$*. Then $V \otimes W$ is semisimple and* $\mathrm{H}^2(\mathfrak{g}, V \otimes W) = 0$. [3]

We also mention a further tool, used in the proofs of Theorems A and B(i), for which we need a definition due to Richardson: Suppose that (G', G) is a pair of reductive algebraic groups such that $G \subseteq G'$ is a closed subgroup. We say that (G', G) is a *reductive pair* provided there is a subspace $\mathfrak{m} \subseteq \mathrm{Lie}(G')$ such that $\mathrm{Lie}(G')$ decomposes as a G-module into a direct sum $\mathrm{Lie}(G') = \mathrm{Lie}(G) + \mathfrak{m}$. Adapting a result from [Her13] we show

Proposition E. *Let (G', G) be a reductive pair and let $H \leq G$ be a closed subgroup scheme. Then if $N_{G'}(H)$ is smooth, $N_G(H)$ is smooth too.*

Acknowledgements

Many thanks to Cornelius Pillen and Brian Parshall who provided a number of helpful hints during the production of this article, and to Alexander Premet for some useful conversations. Also we would like to express our gratitude to the referee for their comments on the paper and encouraging us to elucidate a couple of our arguments along the lines they suggested. The first author acknowledges financial support from ERC Advanced Grant 291512 (awarded to Gunter Malle).

[3] In [Del13], it is proved that if V and W are semisimple modules for an affine group scheme, satisfying $\dim V + \dim W < p+2$, then $V \otimes W$ is semisimple. The semisimplicity part of Theorem D(d) can be deduced from this. In fact, in part of Corollary 8.13 we do prove the semisimplicity statement for arbitrary Lie algebras. Our proof is different to Deligne's, relying just on a theorem of Strade, together with Theorem C.

2 NOTATION AND PRELIMINARIES

Let k be a field of characteristic $p \geq 0$ and let G be an algebraic group defined over k. Unless otherwise noted, k will assumed to be algebraically closed. For all aspects to do with the representation theory of a connected reductive algebraic group G we keep notation compatible with [Jan03]. In particular, R is the root system of G, and h is the associated Coxeter number.

For a closed subgroup $H \leq G$, we consider the scheme-theoretic normaliser $N_G(H)$, respectively centraliser $C_G(H)$ of H in G. We define $N_G(H)$ to be subfunctor of G which takes a k-algebra A and returns the subgroup of elements

$$N_G(H)(A) = \{g \in G(A) : gH(B)g^{-1} = H(B)\}$$

for all A-algebras B. Similarly, the centraliser is defined via

$$C_G(H)(A) = \{g \in G(A) : gh = hg \text{ for all } h \in H(B)\}.$$

Since H is closed, $N_G(H)$ and $C_G(H)$ are closed subgroup schemes of G.

By contrast, for any affine algebraic group H over k, we denote by H_{red} the smooth subgroup with k-points $H_{\mathrm{red}}(k) = H(k)$. As k is algebraically closed, the existence and uniqueness of such a subgroup is explained for example in [Mil12, Prop. 5.1] and (as we will use in the sequel) we have that $N_G(H)_{\mathrm{red}}(k') = N_{G(k')}(H(k'))(k')$ (resp. $C_G(H)_{\mathrm{red}}(k') = C_{G(k')}(H(k'))(k')$) by [Mil12, §VII.6].

Let $\mathfrak{g}$ be a Lie algebra over k. When the characterstic of k is greater than 0, $\mathfrak{g}$ is often referred to as a modular Lie algebra, and as such our reference for the theory is [SF88]. Recall that a Lie algebra $\mathfrak{g}$ is *semisimple* if its solvable radical is zero, and that in characteristic $p > 0$ this is not enough to ensure that it is the direct sum of simple Lie algebras.

Sometimes but not all the time, we will have $\mathfrak{g} = \mathrm{Lie}(G)$ for G an algebraic group, in which cas we refer to $\mathfrak{g}$ as *algebraic*; in this case, $\mathfrak{g}$ will carry the structure of a restricted Lie algebra. Bear in mind that $\mathrm{Lie}(G)$ may not be semisimple even when G is. Examples of this sort only occur in not-very-good characteristic; for instance, $\mathfrak{sl}_2 = \mathrm{Lie}(\mathrm{SL}_2)$ in characteristic 2 gives a restricted structure on the solvable Lie algebra $\mathfrak{sl}_2$ with 1-dimensional centre.

More generally, all restricted Lie algebras are of the form $\mathrm{Lie}(H)$, where H is an infinitesimal group scheme of height one over k. Under this correspondence, the restricted subalgebras of $\mathfrak{g} = \mathrm{Lie}(G)$ correspond to height one subgroup schemes of G. If the centre $Z(\mathfrak{g}) = 0$, then a Lie algebra $\mathfrak{g}$ has at most one restricted structure. In particular, if two semisimple restricted Lie algebras are isomorphic as Lie algebras, they are isomorphic as restricted Lie algebras.

An abelian p-subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ consisting of semsimple elements is called a *torus* of $\mathfrak{g}$. Cartan subalgebras of algebraic Lie algebras are always toral and in fact the Lie algebras of maximal tori of the associated algebraic group. This follows from [Hum67, Thm. 13.3].

If $\mathfrak{g}$ is a restricted Lie algebra, a representation V is called *restricted* provided it is given by a morphism of restricted Lie algebras $\mathfrak{g} \to \mathfrak{gl}(V)$. The following fact follows e.g. from the Kac–Weisfeiler conjecture (see [Pre95, Cor. 3.10]): if G is a simple algebraic group defined in very good characteristic, and if V is an irreducible $\mathfrak{g}$-module with $\dim V < p$, then V is restricted. In particular, it is well-known that V is then obtained by differentiating a simple restricted rational representation of G.

When $\mathfrak{g}$ is a Lie algebra, $\mathrm{Rad}(\mathfrak{g})$ is the solvable radical of $\mathfrak{g}$ and $N(\mathfrak{g})$ is the nilradical of $\mathfrak{g}$. If $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ there is also the radical of V-nilpotent elements $\mathrm{Rad}_V(\mathfrak{g})$. When $\mathfrak{g}$ is restricted, $\mathrm{Rad}_p(\mathfrak{g})$ is the *p-radical* of $\mathfrak{g}$, defined to be the biggest p-nilpotent ideal. Further, $\mathfrak{g}$ is *p-reductive* if the radical $\mathrm{Rad}_p(\mathfrak{g})$ is zero. Recall the following properties from [SF88, §2.1]:

Lemma 2.1. *(a)* $\mathrm{Rad}_p(\mathfrak{g})$ *is contained in the nilradical* $N(\mathfrak{g})$ *and hence in the solvable radical of* $\mathfrak{g}$. *In particular, semisimple Lie algebras are p-reductive.*

(b) $\mathrm{Rad}_p(\mathfrak{g})$ *is the maximal p-nil (that is, consisting of p-nilpotent elements) ideal of* $\mathfrak{g}$.

(c) $\mathfrak{g}/\mathrm{Rad}_p(\mathfrak{g})$ *is p-reductive.*

In particular, by part (b), if $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ is a restricted subalgebra then $\mathrm{Rad}_p(\mathfrak{g}) = \mathrm{Rad}_V(\mathfrak{g})$. If $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ is a restricted Lie subalgebra and G_1 is the height one subgroup scheme of $\mathrm{GL}(V)$ associated to $\mathfrak{g}$, then $\mathfrak{g}$ is p-reductive if and only if G_1 is reductive in the sense that is has no connected normal nontrivial unipotent subgroup schemes. For the usual notion of reductivity of smooth algebraic groups only smooth unipotent subgroups are considered. The relation between these two concepts is as follows:

Proposition 2.2 ([Vas05]). *Let G be a connected reductive algebraic group. Then G has no non-trivial connected normal unipotent subgroup schemes, except if both $p = 2$ and G contains a direct factor isomorphic to* SO_{2n+1} *for some $n \geq 1$.*

Since there are a number of possible definitions, let us be clear on the following: We define a *Borel subalgebra* (resp. *parabolic subalgebra*, resp. *Levi subalgebra*) of $\mathfrak{g}$ to be $\mathrm{Lie}(B)$ (resp. $\mathrm{Lie}(P)$, resp. $\mathrm{Lie}(L)$), where B (resp. P, resp. L) is a Borel (resp. parabolic, resp. Levi subgroup of a parabolic) subgroup of G.

By $P = LQ$ we will denote a parabolic subgroup of G with unipotent radical Q and Levi factor L. We will usually write $\mathfrak{p} = \mathrm{Lie}(P) = \mathfrak{l} + \mathfrak{q}$. A fact that we will use continually during this paper, without proof, is that if H (resp. $\mathfrak{h}$) is a subgroup (resp. subalgebra) of P (resp. $\mathfrak{p}$), such that the projection to the Levi is in a proper parabolic of the Levi, then there is a strictly smaller parabolic $P_1 < P$ (resp. $\mathfrak{p}_1 < \mathfrak{p}$) such that $H \leq P_1$ (resp. $\mathfrak{h} \leq \mathfrak{p}_1$). See [BT65, Prop. 4.4(c)].

We also use the following fact: If $\mathfrak{t} \subseteq \mathfrak{gl}_n$ is a torus, then $C_{\mathrm{GL}_n}(\mathfrak{t})$ is a Levi subgroup (this follows e.g. from the construction of a torus $T \subseteq \mathrm{GL}_n$ in [Die52, Prop. 2] with $C_{\mathrm{GL}_n}(\mathfrak{t}) = C_{\mathrm{GL}_n}(T)$).

Let V be an $\mathfrak{g}$-module and let $\lambda : V \times V \to k$ be a bilinear form on V. We say $\mathfrak{g}$ preserves λ if $\lambda(x(v), w) = -\lambda(v, x(w))$ for all $x \in \mathfrak{g}$, $v, w \in V$.

We recall definitions of the algebraic simple Lie algebras of classical type: those with root systems of types A–D. Then $\mathfrak{o}(V)$ is the set of elements $x \in \mathfrak{gl}(V)$ preserving the form $\lambda(v, w) = v^t w$. $\mathfrak{so}(V)$ is the subset of traceless matrices of $\mathfrak{o}(V)$. On the other hand when $\dim V$ is even, $\mathfrak{sp}(V)$ is the set of elements preserving the form $\lambda(v, w) = v^t Jw$ with $J = [[0, -I_n], [I_n, 0]]$. If char $k \neq 2$ then $\mathfrak{sp}(V)$ and $\mathfrak{so}(V)$ are simple (see below).

We say $\mathfrak{sp}(V)$ is of type C_n with $2n = \dim V$; $\mathfrak{so}(V)$ is of type B_n when $\dim V = 2n + 1$, or type D_n when $\dim V = 2n$. One fact that we shall use often in the sequel is that that for types B–D, parabolic subalgebras are the stabilisers of totally singular subspaces. (See for example, [Kan79].)

Furthermore recall that if G is simple, then $\mathfrak{g}$ is simple at least whenever p is very good. See [Hog82, Cor. 2.7] for a more precise statement. This means in particular that $\mathfrak{sl}(V)$ is simple unless $p|\dim V$, in which case the quotient $\mathfrak{psl}(V) = \mathfrak{sl}(V)/kI$ is simple; we refer to such algebras as type A_n classical Lie algebras, where $\dim V = n+1$. In all cases, we refer to V as the *natural module* for the algebra in question.

We make extensive use of the current state of knowledge of cohomology in this paper, especially in §8. Importantly, recall that the group $\mathrm{Ext}^1_A(V, W)$ (with A either an algebraic group or a Lie algebra) corresponds to the equivalence classes of extensions E of A-modules $0 \to W \to E \to V \to 0$, and that $\mathrm{H}^2(A, V)$ measures the equivalence classes of central extensions B of V by A, equivalence classes of exact sequences $0 \to V \to B \to A \to 0$, where B is either an algebraic group or a Lie algebra. We remind the reader that for restricted Lie algebras, two forms of cohomology are available—the ordinary Lie algebra cohomology, denoted $\mathrm{H}^i(\mathfrak{g}, V)$ or the restricted Lie algebra cohomology (where modules respectively morphisms are assumed to be restricted). Since the latter can always be identified with $\mathrm{H}^i(A, V)$ for A the height one group scheme associated to $\mathfrak{g}$, we shall always use the associated group scheme when we wish to discuss restricted cohomology.

Finally, we record the following theorem of Strade which is a central tool in our study of small-dimensional representations. Let char $k = p > 0$ and let $O_1 = k[X]/X^p$ be the truncated polynomial algebra. Then the first Witt algebra W_1 is the set of derivations of O_1, with basis $\{X^r\partial\}_{0\leq r\leq p-1}$, where ∂ acts on O_1 by differentiation of polynomials. For $p > 2$, W_1 is simple, and for $p > 3$, W_1 is not the Lie algebra of any algebraic group. Since there is a subspace $k \leq O_1$ fixed by W_1, we see that W_1 has a faithful $(p-1)$-dimensional representation for $p > 2$.

THEOREM 2.3 ([Str73, Main theorem]). *Let $\mathfrak{g}$ be a semisimple Lie subalgebra of $\mathfrak{gl}(V)$ over an algebraically closed field k of characteristic $p > 2$ with $p > \dim V$. Then $\mathfrak{g}$ is either a direct sum of algebraic Lie algebras or $p = \dim V + 1$ and $\mathfrak{g}$ is the p-dimensional Witt algebra W_1.*

3 Smoothness of normalisers of reductive subgroups

Let G be a connected reductive algebraic group and let T be a maximal torus in G with associated roots R, coroots $R^\vee$, characters $X(T)$ and cocharacters $Y(T)$. We say that a prime p is *pretty good* for G provided it is good for R and provided that both $X(T)/\mathbb{Z}R$ and $Y(T)/\mathbb{Z}R^\vee$ have no p-torsion. We recall the main result of [Her13].

Proposition 3.1. *Let G be as above, and let $p = \mathrm{char}(k)$. Then p is pretty good for G if and only if all centralisers of closed subgroup schemes in G are smooth.*

Theorem 3.2. *Let G be a connected reductive algebraic group. Then the normalisers $N_G(H)$ of all (smooth) connected reductive subgroups are smooth if p is a pretty good prime for G.*

Proof. Let $H \leq G$ be a closed, connected reductive subgroup of G. We have an exact sequence of group functors

$$1 \to C_G(H) \to N_G(H) \xrightarrow{\mathrm{int}} \mathrm{Aut}(H).$$

Here the first map is the natural inclusion, the second map maps $x \in G$ to the automorphism $\mathrm{int}(x)$ of H given by conjugation with x, and $\mathrm{Aut}(H)$ is the group functor that associates to each k-algebra S the group of automorphisms of the group scheme H_S. By [DGd70, XXIV, Cor. 1.7], we have that $\mathrm{Aut}(H)^0 = \mathrm{int}(H)$ is smooth, which implies that $\mathrm{int}(N_G(H))$ is smooth. By Proposition 3.1, $C_G(H)$ is smooth. Thus the outer terms in the exact sequence of affine group schemes

$$1 \to C_G(H) \to N_G(H) \to \mathrm{int}(N_G(H)) \to 1$$

are smooth, which forces $N_G(H)$ to be smooth. □

Remark 3.3. The implication in the theorem cannot quite be reversed. For example if G is SL_2, $p = 2$ is not pretty good, but a connected reductive subgroup is either trivial, or a torus, whose normaliser is smooth. However, we give examples of non-smooth normalisers of connected reductive subgroups in bad characteristics in Examples 11.6 below.

4 On exponentiation and normalising, and the proof of Theorem B(ii)

Let G be a connected reductive group. We recall the existence of exponential and logarithm maps for p big enough, see [Ser98, Thm. 3] or [Sei00, Prop. 5.2]. We fix a maximal torus T and a Borel subgroup $B = T \ltimes U$ containing T.

THEOREM 4.1. *Assume that $p > h$ ($p \geq h$ for G simply connected), where h is the Coxeter number of G. Then there exists a unique isomorphism of varieties* $\log : G^u \to \mathfrak{g}_{nilp}$, *whose inverse we denote by* $\exp : \mathfrak{g}_{nilp} \to G^u$, *with the following properties:*

(i) $\log \circ \sigma = d\sigma \circ \log$ *for all* $\sigma \in \mathrm{Aut}(G)$*;*

(ii) the restriction of $\log$ *to U is an isomorphism of algebraic groups* $U \to \mathrm{Lie}(U)$*, whose tangent map is the identity; here the group law on* $\mathrm{Lie}(U)$ *is given by the Hausdorff formula;*

(iii) $\log(x_\alpha(a)) = aX_\alpha$ *for every root α and $a \in k$, where* $X_\alpha = dx_\alpha(1)$.

The uniqueness implies that for $G = \mathrm{GL}(V)$, $p \geq \dim V$, exp and log are the usual truncated series.
Recall (cf. [Ser98]) that for a G-module V, the number $n(V)$ is defined as $n(V) = \sup_\lambda n(\lambda)$, where λ ranges over all T-weights of V, and where $n(\lambda) = \sum_{\alpha \in R^+} \langle \lambda, \alpha^\vee \rangle$. For the adjoint module $\mathfrak{g}$, one obtains $n(\mathfrak{g}) = 2h - 2$.

PROPOSITION 4.2. *Let $\rho : G \to \mathrm{GL}(V)$ be a rational representation of G. Suppose that $p > h$ and $p > n(V)$. Let $x \in \mathfrak{g}$ be a nilpotent element. Then*

$$\rho(\exp_G x) = \exp_{\mathrm{GL}}(d\rho(x)).$$

In particular, if $p > 2h - 2$, then $\mathrm{Ad}(\exp_G x) = \exp_{\mathrm{GL}}(\mathrm{ad}(x))$.

Proof. Consider the homomorphism $\varphi : \mathbb{G}_a \to \mathrm{GL}(V)$ given by $\varphi(t) = \rho(\exp_G(t.x))$. Under our assumptions, it follows from [Ser98, Thm. 5] that φ is a morphism of degree $< p$, (i.e. the matrix entries of φ are polynomials of degree less than p in t). Moreover, $d\varphi(1) = d\rho(x)$. By [Ser94, §4], this implies that $d\rho(x)^p = 0$ and that φ agrees with the homomorphism $t \mapsto \exp_{\mathrm{GL}}(t.d\rho(x))$. The claim follows. □

LEMMA 4.3. *Let $X \in \mathfrak{gl}(V)$ be a nilpotent element satisfying $X^n = 0$ for some integer $n \leq p$. Let $l, r \in \mathrm{End}(\mathfrak{gl}(V))$ be left multiplication with X, respectively right multiplication with $-X$. Set $W = W_p(l, r) \in \mathrm{End}(\mathfrak{gl}(V))$, where $W_p(x, y)$ is the the image of $\frac{1}{p}((x + y)^p - x^p - y^p) \in \mathbb{Z}[x, y]$ in $k[x, y]$. Let $\mathfrak{h}$ be a subset of $\mathfrak{gl}(V)$ normalised (resp. centralised) by X. Suppose that $\mathfrak{h} \subseteq \ker(W)$. Then* $\exp(X) \in \mathrm{GL}(V)$ *normalises (resp. centralises) $\mathfrak{h}$.*
In particular, if $p \geq 2n - 1$, then $W = 0$ and so $\exp(X)$ *normalises (resp. centralises) every subspace that is normalised (resp. centralised) by X.*

Proof. Since the nilpotence degree of X is less than p, the exponential $\exp(X) = 1 + X + X^2/2 + \dots$ gives a well-defined element of $\mathrm{GL}(V)$. Moreover, for each $Y \in \mathfrak{h}$ we have the equality

$$\mathrm{Ad}(\exp(X))(Y) = \exp(\mathrm{ad}(X))(Y) = Y + \mathrm{ad}(X)(Y) + \mathrm{ad}(X)^2(Y)/2 + \dots \in \mathfrak{gl}(V).$$

Indeed, we have $\mathrm{ad}(X) = l + r$, and $\mathrm{Ad}(\exp(X)) = \exp(l)\exp(r)$. Now by [Ser94, (4.1.7)], $\exp(l)\exp(r) = \exp(l + r - W)$. Since l and r commute with W, we deduce $(l + r - W)^m(Y) = (l + r)^m(Y)$ for each $m \geq 0$. Thus $\mathrm{Ad}(\exp(X))(Y) = \exp(l + r)(Y) = \exp(\mathrm{ad}(X))(Y)$, as claimed. Hence $\exp(X)$ is contained in $N_{\mathrm{GL}(V)}(\mathfrak{h})$ whenever $X \in \mathfrak{n}_{\mathfrak{gl}(V)}(\mathfrak{h})$ and $\exp(X) \in C_{\mathrm{GL}(V)}(\mathfrak{h})$ whenever $X \in \mathfrak{c}_{\mathfrak{gl}(V)}(\mathfrak{h})$.
Moreover, $W_p(l, r) = \sum_{i=1}^{p-1} c_i l^i r^{p-i}$ for certain non-zero coefficients $c_i \in k$. In particular, this expression vanishes for $p \geq 2n - 1$. □

Corollary 4.4. *Let $\mathfrak{p} = \mathfrak{q} + \mathfrak{l} \subseteq \mathfrak{gl}(V)$ be a parabolic subalgebra, and suppose that $p \geq \dim V$. If $X \in \mathfrak{q}$ normalises a subset $\mathfrak{h} \subseteq \mathfrak{p}$, then so does* $\exp(X)$.

Proof. By Lemma 4.3, it suffices to show that $\mathfrak{p} \subseteq \ker(W)$. Let $0 = V_0 \subseteq V_1 \subseteq \cdots \subseteq V_m = V$ be a flag with the property

$$\mathfrak{p} = \{Y \in \mathfrak{gl}(V) \mid YV_i \subseteq V_i\}$$
$$\mathfrak{q} = \{Y \in \mathfrak{gl}(V) \mid YV_i \subseteq V_{i-1}\}.$$

By assumption, we have $p \geq m$, and therefore all products $X_1 \ldots X_{p+1}$ with all $X_i \in \mathfrak{p}$ and all but one $X_i \in \mathfrak{q}$ vanish on V. In particular $l^i r^{p-i}(Y) = 0$ for all $Y \in \mathfrak{p}$ and hence $W(Y) = 0$. □

Lemma 4.5. *Suppose $\mathfrak{g}$ is a subalgebra of $\mathfrak{gl}(V)$ generated as a k-Lie algebra by a set of nilpotent elements $\{X_i\}$ of nilpotence degree less than p, and let $G = \overline{\langle \exp(t.X_i)\rangle}$ be the closed subgroup of* $\mathrm{GL}(V)$ *generated by* $\exp(t.X_i)$ *for each $t \in k$. Then $\mathfrak{g} \leq$* $\mathrm{Lie}(G)$.

Proof. Since $\mathrm{Lie}(G)$ contains the element $d/dt \exp(t.X_i)|_{t=0}$ it contains each element X_i. Since $\mathfrak{g}$ is generated by the elements X_i, we are done. □

Proof of Theorem B(ii). Let $\mathfrak{h}$ be a subspace of $\mathfrak{g}$ and let $\mathfrak{n} = \mathfrak{n}_{\mathfrak{g}}(\mathfrak{h})$ be the Lie-theoretic normaliser of $\mathfrak{h}$ in $\mathfrak{g}$.
Let $\{x_1, \ldots, x_r\}$ be a set of nilpotent elements generating $\mathfrak{n}$. To show that $N_G(\mathfrak{h})$ is smooth, it suffices to show that each x_i belongs to the Lie algebra of $N_G(\mathfrak{h})_{\mathrm{red}}$.
But for a nilpotent generator x_i, we may consider the smooth closed subgroup $M_i = \overline{\langle \exp(t.x_i) \mid t \in k\rangle}$ of G. By Proposition 4.2, $M_i \subseteq N_G(\mathfrak{h})_{\mathrm{red}}$ and hence $x_i \in \mathrm{Lie}(M_i) \subseteq \mathrm{Lie}(N_G(\mathfrak{h})_{\mathrm{red}})$, as required. □

5 Reductive pairs: Proof of Proposition E

The following definition is due to Richardson [Ric67].

Definition 5.1. Suppose that (G', G) is a pair of reductive algebraic groups such that $G \subseteq G'$ is a closed subgroup. Let $\mathfrak{g}' = \mathrm{Lie}(G)$, $\mathfrak{g} = \mathrm{Lie}(G)$. We say that (G', G) is a *reductive pair* provided there is a subspace $\mathfrak{m} \subseteq \mathfrak{g}'$ such that $\mathfrak{g}'$ decomposes as a G-module into a direct sum $\mathfrak{g}' = \mathfrak{g} \oplus \mathfrak{m}$.

With p sufficiently large, reductive pairs are easy to find.

LEMMA 5.2 ([BHMR11, Thm. 3.1]). *Suppose $p > 2\dim V - 2$ and G is a connected reductive subgroup of* $\mathrm{GL}(V)$. *Then* $(\mathrm{GL}(V), G)$ *is a reductive pair.*

We need a compatibility result for normalisers of subgroup schemes of height one.

LEMMA 5.3. *Let $H \subseteq G$ be a closed subgroup scheme of height one, with $\mathfrak{h} =$* $\mathrm{Lie}(H)$. *Then $N_G(H) = N_G(\mathfrak{h})$ (scheme-theoretic normalisers).*

Proof. We have a commutative diagram

$$\begin{array}{ccc} \mathrm{Hom}(H,H) & \longrightarrow & \mathrm{Hom}_{p-\mathrm{Lie}}(\mathfrak{h},\mathfrak{h}) \\ \downarrow & & \downarrow \\ \mathrm{Hom}(H,G) & \longrightarrow & \mathrm{Hom}_{p-\mathrm{Lie}}(\mathfrak{h},\mathfrak{g}), \end{array}$$

where the horizontal arrows are given by differentiation and are bijective (cf. [DG70, II, §7, Thm. 3.5]). Now if $x \in N_G(\mathfrak{h})$, the map $\mathrm{Ad}(x)_{\mathfrak{h}}$ in the bottom right corner may be lifted via the top right corner to a map in $\mathrm{Hom}(H,H)$. The commutativity of the diagram shows that conjugation by x stabilises H, and hence $x \in N_G(H)$. This works for points x with values in any k-algebra, and hence proves the containment of subgroup schemes $N_G(\mathfrak{h}) \subseteq N_G(H)$. The reverse inclusion is clear. □

We show that the smoothness of normalisers descends along reductive pairs. Let us restate and then prove Propostion E.

PROPOSITION 5.4. *Let (G', G) be a reductive pair and let $H \subseteq G$ be a closed subgroup scheme. If $N_{G'}(H)$ is smooth, then so is $N_G(H)$.*
In particular, if $\mathfrak{h} \subseteq \mathfrak{g}$ is a restricted subalgebra and if $N_{G'}(\mathfrak{h})$ is smooth, then so is $N_G(\mathfrak{h})$.

Proof. The last assertion follows from Lemma 5.3.
Let $H \subseteq G$ be a closed subgroup scheme. We follow the proof of [Her13, Lem. 3.6]. Let $\mathfrak{g}' = \mathfrak{g} \oplus \mathfrak{m}$ be a decomposition of G-modules.
By [DG70, II, §5, Lem. 5.7], we have

$$\begin{aligned} \dim \mathrm{Lie}(N_{G'}(H)) &= \dim \mathfrak{h} + \dim(\mathfrak{g}'/\mathfrak{h})^H = \dim \mathfrak{h} + \dim(\mathfrak{g}/\mathfrak{h})^H + \dim \mathfrak{m}^H \\ &= \dim \mathrm{Lie}(N_G(H)) + \dim \mathfrak{m}^H \geq \dim N_G(H) + \dim \mathfrak{m}^H. \end{aligned}$$

On the left hand side, as $N_{G'}(H)$ is smooth by assumption, we have $\dim N_{G'}(H) = \dim \mathrm{Lie}(N_{G'}(H))$. Thus to show that $N_G(H)$ is smooth, it suffices to show that $\dim N_{G'}(H) - \dim N_G(H) \leq \dim \mathfrak{m}^H$.
Now as in [Her13, Lem. 3.6], one shows that there is a monomorphism of quotient schemes $N_{G'}(H)/N_G(H) \hookrightarrow (G'/G)^H$, and that the tangent space on the right hand side identifies as $T_{\bar{e}}(G'/G)^H \cong \mathfrak{m}^H$. The claim follows. □

6 Lifting of normalising tori and the proof of Theorem A

In this section we let $G = \mathrm{GL}(V)$ and $\mathfrak{h}$ be a subspace of $\mathfrak{g}$. We would like to lift a normaliser $\mathfrak{n}_{\mathfrak{g}}(\mathfrak{h})$ to a subgroup N normalising $\mathfrak{h}$ such that $\mathrm{Lie}(N) = \mathfrak{n}_{\mathfrak{g}}(\mathfrak{h})$. It turns out that the hardest part of this is to find a lift of a maximal torus normalising $\mathfrak{h}$. This is the content of the next lemma.

Lemma 6.1. *Let $G = \mathrm{GL}_n$ with $p > 2^{2n}$ and let $\mathfrak{h} \subseteq \mathfrak{g}$ be any subspace of $\mathfrak{g} = \mathrm{Lie}(G)$. Suppose that $\mathfrak{c} \subseteq \mathfrak{g}$ is a torus normalising $\mathfrak{h}$. Then $\mathfrak{c} = \mathrm{Lie}(C)$ for a torus $C \subseteq N_G(\mathfrak{h})$.*

Proof. Let T be a diagonal maximal torus of GL_n and $\mathfrak{t} = \mathrm{Lie}(T)$. Since $\mathfrak{c}$ consists of semisimple elements, we may assume $\mathfrak{c} \subseteq \mathfrak{t}$.
Since $\mathfrak{c}$ is restricted, it has a basis defined over $\mathbb{F}_p$ of elements $Z_1, \dots, Z_s$ with $Z_i = \mathrm{diag}(z_{i1}, \dots, z_{in})$ and each $z_{ij} \in \mathbb{F}_p$. By [Die52, Prop. 2] we may assume that $\mathfrak{c}$ is a maximal torus of $\mathfrak{n}_{\mathfrak{g}}(\mathfrak{h})$, which we do from now on.
Since k is algebraically closed, we may take a decomposition of $\mathfrak{h}$ into weight spaces for $\mathfrak{c}$. We have $\mathfrak{h} = \mathfrak{h}_0 \oplus \bigoplus_\alpha \mathfrak{h}_\alpha$ where $\mathfrak{h}_0$ is some set of elements commuting with $\mathfrak{c}$, α is a non-trivial linear functional $\mathfrak{c} \to k$ and each $\mathfrak{h}_\alpha$ is a subspace of $\mathfrak{gl}_n$ with $[c, X] = \alpha(c)X$ for $c \in \mathfrak{c}$ and $X \in \mathfrak{h}_\alpha$.
Let $\{X_i\}$ be a basis for $\mathfrak{h}$ with each $X_i \in \mathfrak{h}_0$ or $\mathfrak{h}_\alpha$ for some α as above. Then $\mathfrak{c} = \bigcap_i \mathfrak{n}_{\mathfrak{t}}(\langle X_i \rangle)$. Suppose $c = \mathrm{diag}(c_1, \dots, c_n)$. The condition $c \in \mathfrak{n}_{\mathfrak{t}}(\langle X_i \rangle)$ puts a set of conditions on the c_i. If only one entry of the matrix X_i is non-zero or X_i is diagonal, then $\mathfrak{t}$ normalises X_i, hence the set of conditions is empty. Otherwise, if $(X_i)_{j,k}$ and $(X_i)_{l,m}$ are non-zero, then c normalising $\langle X_i \rangle$ implies $c_j - c_k = c_l - c_m$. Letting $\mathbf{c} = (c_1, \dots, c_n)$ this condition can be rewritten as a linear equation $\mathbf{rc} = 0$, where $\mathbf{r}$ is an appropriate row vector whose entries are all 0, except for up to four, where the non-zero entries take the values, up to signs or permutations, $(1, -1), (2, -2), (1, -2, 1)$ or $(1, -1, -1, 1)$ according to the values of j, k, l and m. The collection of these, say m relations, across i and all pairs of non-zero entries in X_i gives an $m \times n$ integral matrix R so that $c \in \mathfrak{c}$ if and only if it satisfies the equation $R\mathbf{c} = 0$ modulo p. Similarly, if $\chi(t) = \mathrm{diag}(t^{a_1}, \dots, t^{a_n})$ is a cocharacter with image in T, then one checks that $\chi(t)$ normalises $\mathfrak{h}$ if the integral equation $R\mathbf{a} = 0$ where $\mathbf{a} = (a_1, \dots, a_n)$. If the nullity of R is the same modulo p as it is over the integers then for any $c \in \mathfrak{n}_{\mathfrak{t}}(\mathfrak{h})$, there exists a cocharacter χ of $N_T(\mathfrak{h})$ with $d/dt|_{t=1}(\chi(t)) = c$ and we are done. But if the nullity of R modulo p differs from the nullity of R over the integers, then we must have that $p|d_i$ for d_i one of the non-zero elementary divisors of R. Now by the theory of Smith Normal Form, if $r \in \mathbb{N}$ is taken maximal so that there exists a non-vanishing $r \times r$ minor, then the elementary divisors of R are all at most the greatest common divisor of all non-zero $r \times r$ minors. Let M be such an $r \times r$ minor. We are going to argue by induction on r that $|\det(M)| \le 2^{2r}$. Since $r \le n$, the hypothesis will then show that p is not a prime factor of $\det(M)$, as required.
We must have $r \le n$. If there is a row of M containing only elements of modulus 2, then at most 2 of these are non-zero and 2 is a prime factor of

$\det M$; Laplace's formula implies that the remaining matrix has determinant at most $2\det M'$ where M' is a certain $r-1\times r-1$ minor of M, so that we are done by induction. If there are no entries of modulus 2, then each row contains at most 4 entries of modulus 1 and Laplace's formula then implies that $\det M \leq 4\det M'$ where M' is a certain $r-1\times r-1$ minor of M of the required form, so that we are done again by induction. Otherwise there is at least one row with non-zero entries $(1,-2)$ or $(1,-2,1)$. By Laplace's formula and induction, it is now easy to see that $|\det M| \leq 2^{2n-2}+2.2^{2n-2}+2^{2n-2}=2^{2n}$ and we are done. $\square$

We are now in a position to prove Theorem A.

Proof of Theorem A. First consider the case $G = \mathrm{GL}_n$. Let $\mathfrak{h}$ be a subspace of $\mathfrak{g}$ and let $\mathfrak{n} = \mathfrak{n}_{\mathfrak{g}}(\mathfrak{h})$ be the Lie-theoretic normaliser of $\mathfrak{h}$ in $\mathfrak{g}$.
As before, by definition, $\mathfrak{n}$ is a restricted subalgebra of $\mathfrak{g}$. Hence, applying the Jordan decomposition for restricted Lie algebras, we see that $\mathfrak{n}$ is generated by its nilpotent and semisimple elements. Let $\{x_1,\dots,x_r,y_1,\dots,y_s\}$ be such a generating set with $x_1,\dots,x_r$ nilpotent and $y_1,\dots,y_s$ semisimple. To show that $N_G(\mathfrak{h})$ is smooth, it suffices to show that all the elements x_i and y_j belong to the Lie algebra of $N_G(\mathfrak{h})_{\mathrm{red}}$.
For a nilpotent generator x_i, of nilpotence degree at most $n<p$, consider the smooth closed subgroup $M_i = \overline{\langle \exp(t.x_i) \mid t\in k\rangle}$ of G. Since $p>2h-2$, we may apply Proposition 4.2, to obtain $M_i \subseteq N_G(\mathfrak{h})_{\mathrm{red}}$ and hence $x_i \in \mathrm{Lie}(M_i) \subseteq \mathrm{Lie}(N_G(\mathfrak{h})_{\mathrm{red}})$, as required.
It remains to consider the semisimple generators y_i. Let $\mathfrak{t}_i := \langle y_i\rangle_p \leq \mathfrak{n}$ be the torus generated by the p-powers of y_i. By hypothesis, $p>2^{2n}$ and so we may apply Lemma 6.1 to find a torus $T_i \leq N_G(\mathfrak{h})$ such that $\mathrm{Lie}(T_i)=\mathfrak{t}_i$. In particular $y_i \in \mathrm{Lie}(N_G(\mathfrak{h})_{\mathrm{red}})$. This finishes the proof in the case $G=\mathrm{GL}(V)$.
If G is a reductive algebraic group suppose $G\to \mathrm{GL}(V)\cong \mathrm{GL}_d$ is a minimal faithful module for G. Now since $p>2^{2\dim V}$, we have that normalisers of all subspaces of $\mathrm{GL}(V)$ are smooth. But now, by Lemma 5.2, $(\mathrm{GL}(V),G)$ is a reductive pair, so that invoking Proposition 5.4 we obtain that $N_G(\mathfrak{h})$ is smooth. This completes the proof. $\square$

7 Non-semisimple subalgebras of classical Lie algebras. Proof of Theorem C

Suppose $\operatorname{char} k>2$ for this section.
This section provides proofs for some of the claims made in [Ten87]. Here we tackle the proof of Theorem C.

Proposition 7.1 (see [SF88, §5.8, Exercise 1]). *Let $\mathfrak{g}\leq \mathfrak{gl}(V)$ be a Lie algebra acting irreducibly on an $\mathfrak{g}$-module V such that $\mathfrak{g}$ preserves a non-zero bilinear form. Then $\mathfrak{g}$ is semisimple.*

Proof. Assume otherwise. Then $\mathrm{Rad}(\mathfrak{g}) \neq 0$ and we can find an abelian ideal $0 \neq J \lhd \mathfrak{g}$. Take $x \in J$. As $[x^p, y] = \mathrm{ad}(x)^p y \in J^{(1)} = 0$, x^p centralises $\mathfrak{g}$ and we have that $v \mapsto x^p v$ is a $\mathfrak{g}$-homomorphism $V \to V$. Since k is algebraically closed and V is irreducible, Schur's lemma implies that $x^p v = \alpha(x) v$ for some map $\alpha : J \to k$.
Since $\lambda \neq 0$ there are v, w with $\lambda(v, w) = 1$. Now $\alpha(x) = \lambda(x^p v, w) = -\lambda(v, x^p w) = -\alpha(x)$ so $\alpha(x) = 0$. Thus $x^p v = 0$ for all $x \in J$. Hence J acts nilpotently on V and so Engel's theorem gives an element $0 \neq v \in V$ annihilated by J. Since V is irreducible, it follows that $JV = J(\mathfrak{g}v) \leq \mathfrak{g}Jv = 0$. Thus $J = 0$ and $\mathfrak{g}$ is semisimple. □

Since any subalgebra of a classical simple Lie algebra of type B, C or D preserves the associated (non-degenerate) form we get

Corollary 7.2. *If $\mathfrak{h}$ is a non-semisimple subalgebra of a classical simple Lie algebra $\mathfrak{g}$ of type B, C or D then $\mathfrak{h}$ acts reducibly on the natural module V for $\mathfrak{g}$.*

Remark 7.3. If $\mathfrak{g} = \mathfrak{g}_2$ (resp. $\mathfrak{f}_4$, $\mathfrak{e}_7$, $\mathfrak{e}_8$) then a subalgebra acting irreducibly on the self-dual modules V_7 (resp. V_{26}, or V_{25} if $p = 3$, V_{56}, $V_{248} = \mathfrak{e}_8$) is semisimple. Here V_n refers to the usual irreducible module of dimension n.

A subalgebra is *maximal rank* if it is proper and contains a Cartan subalgebra (CSA) of $\mathfrak{g}$. (Note that CSAs of simple algebraic Lie algebras are tori.) Call a subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ an *R-subalgebra* if $\mathfrak{h}$ is contained in a maximal rank subalgebra of $\mathfrak{g}$.
For the following, notice that if $p | \dim V$ then $\mathfrak{sl}(V)$ is not simple, though provided $\mathfrak{sl}(V) \neq \mathfrak{sl}_2$ in characteristic 2, the central quotient $\mathfrak{psl}(V)$ is simple. Now, a subalgebra $\mathfrak{h}$ of $\mathfrak{psl}(V)$ is an R-subalgebra of $\mathfrak{psl}(V)$ if and only if its preimage $\pi^{-1}\mathfrak{h}$ under $\pi : \mathfrak{sl}(V) \to \mathfrak{psl}(V)$ is an R-subalgebra. We say $\mathfrak{h}$ acts reducibly on V if $\pi^{-1}\mathfrak{h}$ does.

Proposition 7.4. *Let $\mathfrak{g}$ be a simple algebraic Lie algebra of classical type and let $\mathfrak{h} \leq \mathfrak{g}$ act reducibly on the natural module V for $\mathfrak{g}$. Then $\mathfrak{h}$ is an R-subalgebra unless $\mathfrak{g} = \mathfrak{so}(V)$ with $\dim V = 2n$ with $\mathfrak{h} \leq \mathfrak{so}(W) \times \mathfrak{so}(W')$ stabilising a decomposition of V into two odd-dimensional, non-degenerate subspaces W and W' of V.*

Proof. Let V be the natural module for $\mathfrak{g}$ and let $W \leq V$ be a minimal $\mathfrak{h}$-submodule, so that $\mathfrak{h} \leq \mathrm{Stab}_{\mathfrak{g}}(W)$. If $\mathfrak{g}$ is of type A then $\mathrm{Stab}_{\mathfrak{g}}(W)$ is $\mathrm{Lie}(P)$ for a (maximal) parabolic P of $\mathrm{SL}(V)$. Hence $\mathfrak{h}$ is an R-subalgebra of $\mathfrak{g}$.
If $\mathfrak{g}$ is of type B, C or D, then consider $U = W \cap W^\perp$; this is the subspace of W whose elements v satisfy $\lambda(v, w) = 0$ for every $w \in W$. Since M preserves λ, this is a submodule of W, hence we have either $U = 0$ or $U = W$ by minimality of W. If the latter, W is totally singular. Thus $\mathrm{Stab}_{\mathfrak{g}} W$ is $\mathrm{Lie}(P)$ for a parabolic subgroup P of the associated algebraic group.
On the other hand, $U = 0$ implies that W is non-degenerate. Then $V = W \oplus W^\perp$ is a direct sum of $\mathfrak{h}$-modules and we see that $\mathrm{Stab}_{\mathfrak{g}} W$ is isomorphic to

(i) $\mathfrak{sp}_{2r} \times \mathfrak{sp}_{2s}$ in case L is of type C, $\dim W = 2s$ and $2r + 2s = \dim V$

(ii) $\mathfrak{so}_r \times \mathfrak{so}_s$ in case L is of type B or D, $\dim W = s$ and $r + s = \dim V$.

Note that by [Bou05, VII, §2, No. 1, Prop. 2] the dimensions of the CSA of a direct product is the sum of the dimensions of the CSAs of the factors. In case (i), the subalgebra described has the $(r+s)$-dimensional CSA arising from the two factors. In case (ii), if $\dim V = 2n+1$ is odd then one of r and s is odd. If r is odd then $\mathfrak{so}_r$ has a CSA of dimension $(r-1)/2$, and $\mathfrak{so}_s$ has a CSA of dimension $s/2$, so that the two together give a CSA of dimension $s/2 + (r-1)/2 = n$. (Similarly if s is odd.) Otherwise $\dim V = 2n$ is even. If $\dim W$ is even then $\mathrm{Stab}_{\mathfrak{g}} W$ contains a CSA of dimension $r/2 + s/2 = n$. If $\dim W$ is odd then we are in the exceptional case described in the proposition. □

Remark 7.5. In the exceptional case, note that $\mathfrak{so}_{2r+1} \times \mathfrak{so}_{2s+1}$ contains a CSA of dimension $r+s$, whereas $\mathfrak{so}_{2n+2}$ contains a CSA of dimension $n+1 = r+s+1$.

COROLLARY 7.6. *Let $\mathfrak{g}$ be of type B, C or D. If $\mathfrak{h}$ is a maximal non-semisimple subalgebra of $\mathfrak{g}$, then $\mathfrak{h}$ is $\mathrm{Lie}(P)$ for P a maximal parabolic of G. In particular, if $\mathfrak{h}$ is any non-semisimple subalgebra of $\mathfrak{g}$, it is an R-subalgebra.*

Proof. Assume otherwise. Then $\mathfrak{h}$ fixes no singular subspace on V. Suppose $\mathfrak{h}$ preserves a decomposition $V = V_1 \perp V_2 \perp \cdots \perp V_n$ on V with n as large as possible, with the V_i all non-degenerate. Then $\mathfrak{h} \leq \mathfrak{g}_1 = \mathfrak{so}(V_1) \times \cdots \times \mathfrak{so}(V_n)$ or $\mathfrak{h} \leq \mathfrak{g}_1 = \mathfrak{sp}(V_1) \times \cdots \times \mathfrak{sp}(V_n)$. Since $\mathfrak{h}$ is non-semisimple, the projection $\mathfrak{h}_1$ of $\mathfrak{h}$ in $\mathfrak{so}(V_1)$ or $\mathfrak{sp}(V_1)$, say, is non-semisimple. Then Proposition 7.1 shows that $\mathfrak{h}$ acts reducibly on V_1. Since $\mathfrak{h}$ stabilises no singular subspace, the proof of Proposition 7.4 shows that $\mathfrak{h}$ stabilises a decomposition of V_1 into two non-degenerate subspaces, a contradiction of the maximality of n. □

Let $\mathfrak{h}$ be a restricted Lie algebra, $I \leq \mathfrak{h}$ an abelian ideal and V an $\mathfrak{h}$-module. Let $\lambda \in I^*$. Recall from [SF88, §5.7] that $\mathfrak{h}^\lambda = \{x \in \mathfrak{h} | \lambda([x,y]) = 0 \text{ for all } y \in I\}$ and $V^\lambda = \{v \in V | x.v = \lambda(x)v \text{ for all } x \in I\}$.

PROPOSITION 7.7. *Let $\mathfrak{h}$ be a non-semisimple subalgebra of $\mathfrak{sl}(V)$ with V irreducible for $\mathfrak{h}$. Then $p | \dim V$.*

Proof. Let $\mathfrak{h}$ be as described and let I be a nonzero abelian ideal of $\mathfrak{h}$. If $\mathfrak{h}_p$ denotes the closure of $\mathfrak{h}$ under the p-mapping, then by [SF88, 2.1.3(2),(4)], I_p is an abelian p-ideal of $\mathfrak{h}_p$. Thus $\mathrm{Rad}\,\mathfrak{h}_p \neq 0$ and $\mathfrak{h}_p$ is non-semisimple. Hence we may assume from the outset that $\mathfrak{h} = \mathfrak{h}_p$ is restricted with nonzero abelian ideal I.

Since $\mathfrak{h}$ acts irreducibly on V, by [SF88, Corollary 5.7.6(2)] there exist $S \in \mathfrak{h}^*$, $\lambda \in I^*$ such that

$$V \cong \mathrm{Ind}_{\mathfrak{h}^\lambda}^{\mathfrak{h}}(V^\lambda, S).$$

If λ is identically 0 on I then V^λ is an $\mathfrak{h}$-submodule. We cannot have $V^\lambda = 0$ (or else $V = 0$) so $V^\lambda = V$ and I acts trivially on V, a contradiction since $I \leq \mathfrak{sl}(V)$.

Hence $\lambda(x) \neq 0$ for some $x \in I$. Suppose $V^\lambda = V$. Then as $x \in \mathfrak{sl}(V)$, we have $\mathrm{tr}_V(x) = \dim V \cdot \lambda(x) = 0$ and thus $p | \dim V$ and we are done. If $\dim V^\lambda < \dim V$, then by [SF88, Prop. 5.6.2] we have $\dim V = p^{\dim L/L^\lambda} \cdot \dim V^\lambda$. Thus again $p | \dim V$, proving the theorem. $\square$

Corollary 7.8. *If $p \nmid \dim V$ then any non-semisimple subalgebra $\mathfrak{h}$ of $\mathfrak{sl}(V)$ acts reducibly on V. Hence it is contained in* $\mathrm{Lie}(P)$ *for P a maximal parabolic of* $\mathrm{SL}(V)$*. In particular $\mathfrak{h}$ is an R-subalgebra.*

Putting together Corollaries 7.6 and 7.8, this completes the proof of Theorem C.

As a first application, the following lemma uses Theorem C to show that p-reductive implies strongly p-reductive. Recall that a restricted Lie algebra is *strongly p-reductive* if it is the direct sum of a central torus and a semisimple ideal.

Lemma 7.9. *Let $\mathfrak{h} \subseteq \mathfrak{gl}_n$ be a subalgebra and let $p > n$. If $\mathfrak{h}$ is p-reductive, it is strongly p-reductive.*

Proof. Take $\mathfrak{p} = \mathfrak{l} + \mathfrak{q}$ a minimal parabolic subalgebra with $\mathfrak{h} \leq \mathfrak{p}$. Set $\mathfrak{h}_l$ to be the image of $\mathfrak{h}$ under the projection $\pi : \mathfrak{p} \to \mathfrak{l}$. Since $p > n$, we have $\mathfrak{l} \cong \mathfrak{gl}(W_1) \times \cdots \times \mathfrak{gl}(W_s) \cong \mathfrak{sl}(W_1) \times \ldots \mathfrak{sl}(W_s) \times \mathfrak{z}$, where $\mathfrak{z}$ is a torus. Let $\mathfrak{s}_i$ be the projection of $\mathfrak{h}_\mathfrak{l}$ to $\mathfrak{sl}(W_i)$, and let $\mathfrak{z}'$ be the projection of $\mathfrak{h}_\mathfrak{l}$ to $\mathfrak{z}$. If the projection of $\mathrm{Rad}(\mathfrak{h}_\mathfrak{l})$ to $\mathfrak{sl}(W_i)$ is non-trivial, then $\mathfrak{s}_i$ is not semisimple. By Theorem C, W_i is not irreducible for $\mathfrak{s}_i$. Thus $\mathfrak{p}$ is not minimal subject to containing $\mathfrak{h}$, a contradiction, proving that all the $\mathfrak{s}_i$ are semisimple. Moreover, $\mathfrak{z}' = Z(\mathfrak{h}_\mathfrak{l})$, as the projection of $\mathfrak{z}$ to each $\mathfrak{sl}(W_i)$ must vanish. This forces $\mathfrak{h}_\mathfrak{l} \subseteq \mathfrak{s}_1 \times \cdots \times \mathfrak{s}_s \times Z(\mathfrak{h}_\mathfrak{l})$ to be strongly p-reductive. As $\mathfrak{h}$ is p-reductive, we have that π is injective on $\mathfrak{h}$, and hence $\mathfrak{h} \cong \mathfrak{h}_l$ is strongly p-reductive. $\square$

8 Complete reducibility and low-degree cohomology for classical Lie algebras: Proof of Theorem D

Let G be a connected reductive algebraic group with root system R and let $G_r \lhd G$ be the rth Frobenius kernel for any $r \geq 1$. It is well-known that the representation theory of G_1 and $\mathfrak{g}$ are very closely related. In this section we recall results on the cohomology of small G_r-modules and use a number of results of Bendel, Nakano and Pillen to prove that small G_r-modules are completely reducible with essentially one class of exceptions. We do this by examining $\mathrm{Ext}^1_{G_r}(L(\lambda), L(\mu))$ for two simple modules $L(\lambda)$ and $L(\mu)$ of bounded dimension or weight. While we are at it, we also get information about $\mathrm{H}^2(G_1, L(\lambda))$. In a further subsection, we then go on to use this to prove the analogous statements for $\mathfrak{g}$-modules. One crucial difference we notice is with central extensions: $\mathrm{H}^2(\mathfrak{g}, k)$ tends to be zero, whereas $\mathrm{H}^2(G_1, k)$ is almost always not; c.f. Corollary 8.2 and Theorem 8.9.

All the notation in this section is as in [Jan03, List of Notations, p. 569]: In particular, for a fixed maximal torus $T \leq G$, we denote by R the corresponding

root system, by R^+ a choice of positive roots with corresponding simple roots $S \subseteq R^+$, by $X(T)_+ \subseteq X(T)$ the dominant weights inside the character lattice, by $L(\lambda)$ the simple G-module of highest weight $\lambda \in X(T)_+$, by $H^0(\lambda)$ the module induced from λ with socle $L(\lambda)$, by $C_{\mathbb{Z}}$ (resp. $\bar{C}_{\mathbb{Z}}$) the dominant weights inside the lowest alcove (respectively, in the closure of the lowest alcove). If G is simply connected, we write $\omega_i \in X(T)_+$ for the fundamental dominant weight corresponding to $\alpha_i \in S = \{\alpha_1, \ldots, \alpha_l\}$.
Let us recall some results from [McN02] which show the interplay between the conditions that, relative to p, (i) modules are of small dimension; (ii) their high weights are small; and (iii) the Coxeter number is small.

PROPOSITION 8.1 ([McN02, Prop. 5.1]). *Let G be simple and simply connected, let L be a simple non-trivial restricted G-module with highest weight $\lambda \in X(T)_+$ and suppose that* $\dim L \leq p$. *Then*

(i) We have $\lambda \in \bar{C}_{\mathbb{Z}}$.

(ii) We have $\lambda \in C_{\mathbb{Z}}$ if and only if $\dim L < p$.

(iii) We have $p \geq h$. If moreover $\dim L < p$ *then $p > h$.*

(iv) If R is not of type A and $\dim L = p$ *then $p > h$. If $p = h$ and* $\dim L = p$ *then $R = A_{p-1}$ and $\lambda = \omega_i$ with $i \in \{1, p-1\}$.*

8.1 COHOMOLOGY AND COMPLETE REDUCIBILITY FOR SMALL G_1-MODULES

We need values of $\mathrm{H}^i(G_1, H^0(\mu))$ for $\mu \in \bar{C}_{\mathbb{Z}}$ and $i = 1$ or 2. Thus $H^0(\mu) = L(\mu)$.

PROPOSITION 8.2. *Let G be simple and simply connected and suppose $L = L(\mu)$ with $\mu \in \bar{C}_{\mathbb{Z}}$ and $p \geq 3$. Then:*
(i) we have $\mathrm{H}^1(G_1, L) = 0$ unless G is of type A_1, $L = L(p-2)$ and in that case $\mathrm{H}^1(G_1, L)^{[-1]} \cong L(1)$;
(ii) suppose $p > h$. Then we have $\mathrm{H}^2(G_1, L) = 0$ unless: $L = k$ and $\mathrm{H}^2(G_1, k)^{[-1]} \cong \mathfrak{g}^*$; or $G = \mathrm{SL}_3$, with $\mathrm{H}^2(G_1, L(p-3, 0))^{[-1]} \cong L(0, 1)$ and $\mathrm{H}^2(G_1, L(0, p-3))^{[-1]} \cong L(1, 0)$.

Proof. Part (i) is immediate from [BNP02, Corollary 5.4 B(i)]. The A_1 result is well known. Part (ii) requires some argument. If $\mathrm{H}^2(G_1, H^0(\mu)) \neq 0$ then since $p > h$ we may assume $\mu \in C_{\mathbb{Z}}$. Now, the values of $\mathrm{H}^2(G_1, H^0(\mu))^{[-1]}$ are known from [BNP07, Theorem 6.2]. It suffices to find those that are non-zero for which $\mu \in C_{\mathbb{Z}} \setminus \{0\}$. All of these have the form $\mu = w.0 + p\lambda$ for $l(w) = 2$ and $\lambda \in X(T)_+$. Now, if $l(w) = 2$, we have $-w.0 = \alpha + \beta$ for two distinct roots $\alpha, \beta \in R^+$ (cf. [BNP07, p. 166]). To have $w.0 + p\lambda$ in the lowest alcove, one needs $\langle w.0 + p\lambda + \rho, \alpha_0^\vee \rangle < p$. Now $\langle p\lambda, \alpha_0^\vee \rangle \geq p$ so $\langle w.0 + \rho, \alpha_0^\vee \rangle < 0$. Thus $m := \langle \alpha + \beta, \alpha_0^\vee \rangle > h - 1$. Now one simply considers the various cases. If G is simply-laced, then the biggest value of $\langle \alpha, \alpha_0^\vee \rangle$ is 2, when $\alpha = \alpha_0$ and 1 otherwise, thus $m > h - 1$ implies $h \leq 3$. Thus we get $G = \mathrm{SL}_3$, and this case is calculated in [Ste12, Prop. 2.5]. If $G = G_2$ we have m at most 5, giving h at most 5, a contradiction. If G is type B, C or F, then m is at most 4, so $G = \mathrm{Sp}_4$, $p \geq 5$ and this is calculated in [Ibr12, Prop. 4.1]. One checks that all μ such that $\mathrm{H}^2(G_1, L(\mu)) \neq 0$ have $\mu \notin C_{\mathbb{Z}}$. □

Remark 8.3. All the values of $\mathrm{H}^2(G_r, H^0(\lambda))^{[-1]}$ are known for all λ by [BNP07, Theorem 6.2] ($p \geq 3$) and [Wri11] ($p = 2$). For example, $\mathrm{H}^2(G_1, k)^{[-1]} \cong \mathfrak{g}^*$ also when G is of type A_1 and $p = 2$. Even for $\lambda = 0$ there are quite a few exceptional cases when $p = 2$: see [Wri11, C.1.4]. There are also two exceptional cases for $p = 3$, for A_2 and G_2, see [BNP07, Theorem 6.2].

One can go further in the case of 1-cohomology to include extensions between simple modules:

Lemma 8.4 ([BNP02, Corollary 5.4 B(i)]). *Let G be a simple, simply connected algebraic group not of type A_1. If $p > 2$ then $\mathrm{Ext}^1_{G_r}(L(\lambda), L(\mu)) = 0$ for all $\lambda, \mu \in \bar{C}_{\mathbb{Z}}$.*

We will use the above result to show that small G_r-modules are completely reducible, but we must first slightly soup it up before we use it.

Lemma 8.5. *Let G be a simple, simply connected algebraic group not of type A_1 and $p > 2$.*
(i) We have $\mathrm{Ext}^1_{G_r}(L(\lambda)^{[s]}, L(\mu)^{[t]}) = 0$ for all $\lambda, \mu \in \bar{C}_{\mathbb{Z}}$ and $s, t \geq 0$.
(ii) For $\lambda, \mu \in X_r(T)$, let $\lambda = \lambda_0 + p\lambda_1 + \cdots + p^{r-1}\lambda_{r-1}$ and $\mu = \mu_0 + p\mu_1 + \cdots + p^{r-1}\mu_{r-1}$ be their p-adic expansions. Suppose we have $\lambda_i, \mu_i \in \bar{C}_{\mathbb{Z}}$ for each i. Then $\mathrm{Ext}^1_{G_r}(L(\lambda), L(\mu)) = 0$.

Proof. (i) Clearly we may assume $s, t < r$. When $r = 1$ the result follows from Lemma 8.4. So assume $r > 1$. Without loss of generality (dualising if necessary) we may assume $s \leq t$. Suppose $s > 0$ and consider the following subsequence of the five-term exact sequence of the LHS spectral sequence applied to $G_s \lhd G_r$ (see [Jan03, I.6.10]):

$$\begin{aligned}0 \to \mathrm{Ext}^1_{G_{r-s}}(L(\lambda), L(\mu)^{[t-s]}) &\to \mathrm{Ext}^1_{G_r}(L(\lambda)^{[s]}, L(\mu)^{[t]}) \\ &\to \mathrm{Hom}_{G_{r-s}}(L(\lambda), \mathrm{Ext}^1_{G_s}(k,k)^{[-s]} \otimes L(\mu)^{[t-s]}) \to 0.\end{aligned}$$

Since $\mathrm{Ext}^1_{G_s}(k,k) = 0$, we have

$$\mathrm{Ext}^1_{G_{r-s}}(L(\lambda), L(\mu)^{[t-s]}) \cong \mathrm{Ext}^1_{G_r}(L(\lambda)^{[s]}, L(\mu)^{[t]}),$$

and the left-hand side vanishes by induction, so we may assume $s = 0$. There is another exact sequence

$$\begin{aligned}0 \to \mathrm{Ext}^1_{G_{r-1}}(k, \mathrm{Hom}_{G_1}(L(\lambda), L(\mu)^{[t]})^{[-1]}) &\to \mathrm{Ext}^1_{G_r}(L(\lambda), L(\mu)^{[t]}) \\ &\to \mathrm{Hom}_{G_{r-1}}(k, \mathrm{Ext}^1_{G_1}(L(\lambda), L(\mu)^{[t]})^{[-1]}) = 0,\end{aligned}$$

where the last term vanishes by induction. If $t = 0$ then as $\lambda \neq \mu$, the first term of the sequence vanishes and we are done. So we may assume $t > 0$. Now we can rewrite the first term as $\mathrm{Ext}^1_{G_{r-1}}(k, \mathrm{Hom}_{G_1}(L(\lambda), k)^{[-1]} \otimes L(\mu)^{[t-1]})$. If this expression is non-trivial we have $\lambda = 0$ and $\mathrm{Ext}^1_{G_{r-1}}(k, L(\mu)^{[t-1]})$ vanishes by induction, which completes the proof.
(ii) Suppose i is the first time either λ_{i-1} or μ_{i-1} is non-zero. Without loss of generality, $\lambda_{i-1} \neq 0$. Write $\lambda = \lambda^i + p^i \lambda'$ and take a similar expression for μ. Then there is an exact sequence

$$0 \to \mathrm{Ext}^1_{G_{r-i}}(L(\lambda'), \mathrm{Hom}_{G_i}(L(\lambda^i), L(\mu^i))^{[-i]} \otimes L(\mu')) \to \mathrm{Ext}^1_{G_r}(L(\lambda), L(\mu))$$
$$\to \mathrm{Hom}_{G_{r-i}}(L(\lambda'), \mathrm{Ext}^1_{G_i}(L(\lambda^i), L(\mu^i))^{[-i]} \otimes L(\mu')).$$

We have $L(\lambda^i) = L(\lambda_{i-1})^{[i-1]}$ and $L(\mu^i) = L(\mu_{i-1})^{[i-1]}$. Hence the right-hand term vanishes by part (i). The left-hand term is non-zero only if $\lambda^i = \mu^i$ and then we get $\mathrm{Ext}^1_{G_{r-i}}(L(\lambda'), L(\mu')) \cong \mathrm{Ext}^1_{G_r}(L(\lambda), L(\mu))$. Thus the result follows by induction on r. ☐

We put these results together to arrive at an analogue of Jantzen's well-known result [Jan97] that G-modules for which $\dim V \leq p$ are completely reducible.

Proposition 8.6. *Let G be a simple, simply connected algebraic group and let $\dim V \leq p$ be a G_r-module. Then exactly one of the following holds:*
(i) V is a semisimple G_r-module;
(ii) G is of type A_1, $p > 2$, $r = 1$, $\dim V = p$ and V is uniserial, with composition factors $L(p-2-s)$ and $L(s)$ with $0 \leq s \leq p-2$;
(iii) G is of type C_n with $n \geq 1$, $p = 2$ and V is uniserial with two trivial composition factors.

Proof. Assume V has only trivial composition factors. We have $\mathrm{Ext}^1_{G_r}(k,k) \neq 0$ if and only if $p = 2$ and G is of type C_n, in which case $\mathrm{Ext}^1_{G_r}(k,k)^{[-r]} \cong L(\omega_1)$; [Jan03, II.12.2]. This is case (iii).
Otherwise, $p > 2$ and $\mathrm{Ext}^1_{G_r}(L(\lambda), L(\lambda)) = 0$ for all $\lambda \in X_r(T)$ by [Jan03, II.12.9].
Assume G is not of type A_1. By assumption, V has a non-trivial composition factor with $\dim V \leq p$. Then $p > 2$ and the hypotheses of Lemma 8.4 hold. Since $\dim V \leq p$, by Proposition 8.1 any G_r-composition factor $L(\lambda)$ of V has a p-adic expansion $\lambda = \lambda_0 + \cdots + p^{r-1}\lambda_r$ with each $\lambda_i \in \bar{C}_{\mathbb{Z}}$. If there were a non-split extension $0 \to L(\lambda) \to V \to V/L(\lambda) \to 0$ then there would be a non-split extension of $L(\lambda)$ by $L(\mu)$ for $L(\mu)$ a composition factor of V, also of the same form. But by Lemma 8.5(ii) we have $\mathrm{Ext}^1_{G_r}(L(\lambda), L(\mu)) = 0$, hence this is impossible and $L(\lambda)$ splits off as a direct summand. Induction on the direct complement completes the proof in this case.
If G is of type A_1 then the G_r-extensions of simple modules are well known. If $r > 1$ with $\lambda, \mu \in X_r(T)$ then $\dim \mathrm{Ext}^1_{G_r}(L(\lambda), L(\mu)) = \dim \mathrm{Ext}^1_G(L(\lambda), L(\mu))$ and this must vanish whenever $\dim L(\lambda) + \dim L(\mu) \leq p$. If $r = 1$, then the only pairs of G_1-linked weights are s and $p-2-s$ with $\mathrm{Ext}^1_{G_1}(L(s), L(p-2-s)) \cong$

$L(1)^{[1]}$ as G-modules. Here we have $\dim L(s) + \dim L(p-s-2) = p$ giving case (ii). □

The following two corollaries are immediate, in the first case, the passage from G being simple to being reductive is trivial (consider the cover of G by the product of the radical and the simply connected cover of the derived group).

Corollary 8.7. *Let G be a connected reductive algebraic group and let V be a G_r-module with $p > \dim V$. Then V is semisimple.*

Corollary 8.8. *Let G be connected reductive and $G_r \leq \mathrm{GL}(V)$ with $\dim V \leq p$. Then either G_r is completely reducible on V or $\dim V = p$, G is of type A_1, $r = 1$ and G_r is in a maximal parabolic of $\mathrm{GL}(V)$ acting indecomposably on V as described in case (ii) of Proposition 8.6.*
Moreover, if $\mathfrak{g}$ is a p-reductive subalgebra of $\mathrm{GL}(V)$ with $\dim V < p$ then $\mathfrak{g}$ acts semisimply on V.

Proof. If G is not simple, it can be written as HK with H and K non-trivial mutually centralising connected reductive subgroups with maximal tori S and T say. The Frobenius kernels $H_1, K_1 \leq G_1 \leq G_r$ are also mutually centralising, so that H_1 is in the centraliser of T_1. Now the centraliser of T_1 is a proper Levi subgroup of $\mathrm{GL}(V)$, hence restriction of V to H_r has at least one trivial direct factor, with direct complement W say, $\dim W < p$. Thus by Corollary 8.7, W is completely reducible for H_r and by symmetry, for K_r. Thus W is completely reducible for $K_rH_r = G_r$.
Otherwise, G is simple and Proposition 8.6 gives the result (note that case (iii) does not occur due to dimension restrictions).
For the last part, Lemma 7.9 implies that $\mathfrak{g}$ is the direct sum of a semisimple ideal and a torus, and we may hence assume that $\mathfrak{g}$ is a semisimple restricted subalgebra of $\mathfrak{gl}(V)$. If $\mathfrak{g}$ is not irreducible on V, then by Theorem 2.3 there exists a semisimple group G with $\mathfrak{g} = \mathrm{Lie}(G)$. Now the result follows from the case G_1 above. □

8.2 Cohomology and complete reducibility for small $\mathfrak{g}$-modules

We now transfer our results to the ordinary Lie algebra cohomology for $\mathfrak{g}$. Recall the exact sequence [Jan03, I.9.19(1)]:

$$\begin{aligned} 0 &\to \mathrm{H}^1(G_1, L) \to \mathrm{H}^1(\mathfrak{g}, L) \to \mathrm{Hom}^s(\mathfrak{g}, L^{\mathfrak{g}}) \\ &\to \mathrm{H}^2(G_1, L) \to \mathrm{H}^2(\mathfrak{g}, L) \to \mathrm{Hom}^s(\mathfrak{g}, \mathrm{H}^1(\mathfrak{g}, L)) \end{aligned} \tag{1}$$

The following theorem is the major result of this section.

Theorem 8.9. *Let $\mathfrak{g} = \mathrm{Lie}(G)$ be semisimple. Then:*

(a) If $p > h$ with $\mu \in \bar{C}_{\mathbb{Z}}$ then either $\mathrm{H}^2(\mathfrak{g}, L(\mu)) = 0$, or one of the following holds: (i) $\mathfrak{g}$ contains a factor $\mathfrak{sl}_3$ and $L(\mu)$ contains a tensor factor of $L(p-3, 0)$ or $L(0, p-3)$ for this $\mathfrak{sl}_3$; (ii) $\mathfrak{g}$ contains a factor $\mathfrak{sl}_2$ and $L(\mu)$ has a tensor factor $L(p-2)$ for this $\mathfrak{sl}_2$.

(b) If $p > 2$ is very good for G then $\mathrm{H}^2(\mathfrak{g}, k) = 0$.

(c) If $p > 2$ is very good for G and $\lambda, \mu \in \bar{C}_{\mathbb{Z}}$ we have $\mathrm{Ext}^1_{\mathfrak{g}}(L(\lambda), L(\mu)) = 0$, *or G contains a factor of type A_1, $L(\lambda)$ and $L(\mu)$ are simple modules for that factor, $\lambda = s < p-1$, $\mu = p-2-s$ and we have* $\mathrm{Ext}^1_{\mathfrak{g}}(L(\lambda), L(\mu))^{[-1]} \cong L(1)$.

Proof. We may assume that G is simply connected, since the condition on p implies that $\mathfrak{g} = \mathfrak{g}_1 \times \mathfrak{g}_2 \cdots \times \mathfrak{g}_s$. Now one can reduce to the case that G is simple using a Künneth formula. To begin with, any simple module $L(\lambda)$ for $\mathfrak{g} = \mathfrak{g}_1 \times \mathfrak{g}_2 \times \cdots \times \mathfrak{g}_s$ is a tensor product of simple modules $L(\lambda_1) \otimes \cdots \otimes L(\lambda_s)$ for the factors. Then by the Künneth formula $\dim \mathrm{Ext}^1_{\mathfrak{g}}(L(\lambda), L(\mu)) \neq 0$ implies that $\lambda_i = \mu_i$ for all $i \neq j$, some $1 \leq j \leq s$ and $\mathrm{Ext}^1_{\mathfrak{g}}(L(\lambda), L(\mu)) \cong \mathrm{Ext}^1_{\mathfrak{g}_j}(L(\lambda_j), L(\mu_j))$. This means we may assume G to be simple in (c). For $\mathrm{H}^2(\mathfrak{g}, L(\lambda))$ to be non-zero one must have all $\lambda_i = 0$ for all $i \neq j, k$ some $1 \leq j < k \leq s$ and then

$$\begin{aligned}\mathrm{H}^2(\mathfrak{g}, L(\lambda)) = \mathrm{H}^2(\mathfrak{g}_j, L(\lambda_j)) \otimes \mathrm{H}^0(\mathfrak{g}_k, L(\lambda_k)) \oplus \mathrm{H}^1(\mathfrak{g}_j, L(\lambda_j)) \otimes \mathrm{H}^1(\mathfrak{g}_k, L(\lambda_k)) \\ \oplus \mathrm{H}^0(\mathfrak{g}_j, L(\lambda_j)) \otimes \mathrm{H}^2(\mathfrak{g}_k, L(\lambda_k)).\end{aligned}$$

Now first suppose that both λ_j and λ_k are non-trivial. Then only the second direct summand in $\mathrm{H}^2(\mathfrak{g}, L(\lambda))$ survives, and by (1) it coincides with the tensor product of the 1-cohomology groups of the corresponding Frobenius kernels. By Proposition 8.2, non-vanishing would force $\lambda_j = p-2 = \lambda_k$ and $\mathfrak{g}_j = \mathfrak{g}_k = \mathfrak{sl}_2$ giving one exceptional case.

Next we treat the case $\lambda_k = 0$ and λ_j non-trivial. Again by (1) and Proposition 8.2, we obtain $\mathrm{H}^2(\mathfrak{g}, L(\lambda)) = \mathrm{H}^2(\mathfrak{g}_j, L(\lambda_j))$, and we are in the case where G is simple and $L(\lambda)$ non-trivial. In case $\mathfrak{g} = \mathfrak{sl}_2$, the result follows from [Dzh92]. So suppose $\mathfrak{g} \neq \mathfrak{sl}_2$. Setting $L = L(\mu)$ in (1) we see that if $\mu \neq 0$ we have $\mathrm{H}^1(\mathfrak{g}, L) \cong \mathrm{H}^1(G_1, L)$ and the right-hand side is zero by Lemma 8.4. Thus we also have $\mathrm{H}^2(\mathfrak{g}, L) \cong \mathrm{H}^2(G_1, L)$ and the latter is zero by Proposition 8.2 unless $\mathfrak{g} = \mathfrak{sl}_3$ and the exception is as in the statement of the Theorem, since we have excluded the A_1 case.

Finally, the case $\lambda_j = \lambda_k = 0$ reduces by the above to the case G simple, $L = k$ and the claim that $\mathrm{H}^2(\mathfrak{g}, k) = 0$. Here we have $\mathrm{H}^1(\mathfrak{g}, k) \cong (\mathfrak{g}/[\mathfrak{g}, \mathfrak{g}])^*$ and this is zero since p is very good and $\mathfrak{g}$ is semisimple. We also have $\mathrm{H}^2(G_1, k)^{[-1]} \cong \mathfrak{g}^*$. The injective map $\mathrm{Hom}^s(\mathfrak{g}, L^{\mathfrak{g}}) \to \mathrm{H}^2(G_1, L)$ is hence an isomorphism, which forces $\mathrm{H}^2(\mathfrak{g}, k) = 0$ in the sequence (1). This also proves (b).

Now we prove the statement (c) under the assumption that G is simple. We have an isomorphism $\mathrm{Ext}^1_{\mathfrak{g}}(L(\lambda), L(\mu)) \cong \mathrm{H}^1(\mathfrak{g}, L(\mu) \otimes L(\lambda)^*)$. Let $M = L(\mu) \otimes L(\lambda)^*$. If $\lambda \neq \mu$, then applying the exact sequence (1) to M yields $\mathrm{H}^1(\mathfrak{g}, M) \cong \mathrm{H}^1(G_1, M)$ and the latter is zero by Lemma 8.4 if G is not of type A_1 and well-known if G is of type A_1. Hence we may assume $\lambda = \mu$. The assignation of L to the sequence (1) is functorial, thus, associated to the G-map $k \to M \cong$

$\mathrm{Hom}_k(L, L)$, there is a commutative diagram

$$\begin{tikzcd} 0 \arrow[r] & \mathrm{H}^1(\mathfrak{g}, k) = 0 \arrow[r] \arrow[d] & \mathrm{Hom}^s(\mathfrak{g}, k^{\mathfrak{g}}) \cong (\mathfrak{g}^*)^{[1]} \arrow[r, "\cong"] \arrow[d, "\cong"'] & \mathrm{H}^2(G_1, k) \arrow[d, "\theta"'] \\ 0 \arrow[r] & \mathrm{H}^1(\mathfrak{g}, M) \arrow[r] & \mathrm{Hom}^s(\mathfrak{g}, M^{\mathfrak{g}}) \cong (\mathfrak{g}^*)^{[1]} \arrow[r, "\zeta"] & \mathrm{H}^2(G_1, M) \end{tikzcd},$$

where the natural isomorphism $k^{\mathfrak{g}} \to M^{\mathfrak{g}}$ induces the middle isomorphism and the top right isomorphism has been discussed already. We want to show that ζ is injective, since then it would follow that $\mathrm{H}^1(\mathfrak{g}, M) = 0$. To do this it suffices to show that θ is an injection $(\mathfrak{g}^*)^{[1]} \to \mathrm{H}^2(G_1, M)$ and for this, it suffices to show that the simple G-module $(\mathfrak{g}^*)^{[1]}$ does not appear as a submodule of $\mathrm{H}^1(G_1, M/k)$. Now since $\lambda \in \bar{C}_{\mathbb{Z}}$ we have $L(\lambda) \cong H^0(\lambda)$ and so by [Jan03, II.4.21], M has a good filtration. The socle of any module $H^0(\mu)$ with $\mu \in X^+$ is simple. Thus the submodule $k \leq M$ constitutes a section of this good filtration, with M/k also having a good filtration.

The G-modules $\mathrm{H}^1(G_1, H^0(\mu))$ have been well-studied by Jantzen [Jan91] and others. In order to have $(\mathfrak{g}^*)^{[1]}$ a composition factor of $\mathrm{H}^1(G_1, H^0(\mu))$, we would need $\mathfrak{g} \cong \mathfrak{g}^* \cong H^0(\omega_\alpha)$ where $\mu = p\omega_\alpha - \alpha$ and α is a simple root with ω the corresponding fundamental dominant weight; [BNP04, Theorem 3.1(A,B)]. Now for type A_n, with $p \nmid n+1$, we have $\mathfrak{g} = L(2\omega_1)$ if $n = 1$ and $\mathfrak{g} = L(\omega_1 + \omega_n)$ else; and for type B_2, we have $\mathfrak{g} = L(2\omega_2)$, ruling these cases out. For the remaining types, we have

Type	B_n,C_n	D_n	E_6	E_7	E_8	F_4	G_2
$\mathfrak{g} \cong L(\omega_\alpha)$ for $\omega_\alpha =$	ω_2	ω_2	ω_2	ω_1	ω_8	ω_1	ω_2
$\langle p\omega_\alpha - \alpha, \alpha_0^\vee \rangle$	$2p$	$2p$	$2p-1$	$2p-1$	$2p-1$	$2p$	$3p$

On the other hand, since $\lambda \in \bar{C}_{\mathbb{Z}}$ it satisfies $\langle \lambda + \rho, \alpha_0^\vee \rangle \leq p$, i.e. $\langle \lambda, \alpha_0^\vee \rangle \leq p - h + 1$. Hence any high weight μ of $M = L \otimes L^*$ satisfies $\langle \mu, \alpha_0^\vee \rangle \leq 2p - 2h + 2$. Looking at the above table, it is easily seen that this is a contradiction. Thus $(\mathfrak{g}^*)^{[1]}$ is not a composition factor of $\mathrm{H}^1(G_1, M/k)$ and the result follows. $\square$

Remarks 8.10. (i) When $\lambda \neq \mu$ in the proof of the above proposition, one also sees that there is an isomorphism $\mathrm{Ext}^2_{G_1}(L(\lambda), L(\mu)) \cong \mathrm{Ext}^2_{\mathfrak{g}}(L(\lambda), L(\mu))$ but we do not use this fact in the sequel.

(ii) The conclusion of the theorem is incorrect if G is reductive but not semisimple. For example, if G is a torus, then $\mathfrak{g}$ is an abelian Lie algebra, and $\mathrm{H}^1(\mathfrak{g}, k)$ is non-trivial. For instance the two-dimensional non-abelian Lie algebra is a non-direct extension of k by k. One also has $\mathrm{H}^2(k \times k, k) \neq 0$ by the Künneth formula: for example the Heisenberg Lie algebra is a non-split extension of k by $k \times k$.

(iii) When $p = 3$ and $G = \mathrm{SL}_3$, then $H^2(G_1, k)^{[-1]} \cong \mathfrak{g}^* \oplus L(\omega_1) \oplus L(\omega_2)$, by [BNP07, Theorem 6.2]. Thus the same argument shows that $\mathrm{H}^2(\mathfrak{g}, k) \cong L(\omega_1) \oplus L(\omega_2)$. It follows from the Künneth formula that if G is a direct product of n copies of SL_3 then $\mathrm{H}^2(\mathfrak{g}, k) \cong [L(\omega_1) \oplus L(\omega_2)]^{\oplus n}$.

(iv) In part (a) of the theorem, one can be more specific. If $\mathfrak{g} = \mathfrak{sl}_2$ then [Dzh92] shows that $\mathrm{H}^2(\mathfrak{g}, L(p-2))$ is isomorphic to $L(1)^{[1]}$ as a G-module. If $\mathfrak{g} = \underbrace{\mathfrak{sl}_2 \times \cdots \times \mathfrak{sl}_2}_{n \text{ times}} \times \mathfrak{h}$ then one can show moreover that $\mathrm{H}^2(\mathfrak{g}, L(\mu))$ is non-zero only if

$$L(\mu) \cong L(\mu_1) \otimes \cdots \otimes L(\mu_n) \otimes L(\mu_{n+1})$$

with each $\mu_i \in \{0, p-2\}$ and $\mu_{n+1} = 0$. Let r be the number of times $\mu_i = p-2$. Then, the Künneth formula shows that

$$\dim \mathrm{H}^2(\mathfrak{g}, L(\mu)) = \begin{cases} 0 \text{ if } r = 0; \\ 2 \text{ if } r = 1; \\ 4 \text{ if } r = 2; \\ 0 \text{ otherwise.} \end{cases}$$

We use the theorem above to get analogues of Corollary 8.8 for Lie algebra representations.

PROPOSITION 8.11. *Let G be a simple algebraic group with $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$ and let $\dim V \leq p$ be a $\mathfrak{g}$-module. Then exactly one of the following holds:*
(i) V is a semisimple $\mathfrak{g}$-module;
(ii) G is of type A_1, $\dim V = p$ and V is uniserial, with composition factors $L(p-2-s)$ and $L(s)$.

Proof. The proof is similar to Proposition 8.6. Since $\dim V \leq p$, any composition factor of V is a restricted simple $\mathfrak{g}$-module, or V is simple. Since $\mathrm{Ext}^1_{\mathfrak{g}}(k, k) = \mathrm{H}^1(\mathfrak{g}, k) \cong (\mathfrak{g}/[\mathfrak{g}, \mathfrak{g}])^* = 0$, if V consists only of trivial composition factors then V is semisimple. Thus we may assume that $\mathfrak{g}$ contains a non-trivial composition factor L. Then either $\dim L = p$ and V is simple, or $p > h$ by Lemma 8.1(iii). By the condition on V, any two distinct composition factors, $L(\lambda)$ and $L(\mu)$ satisfy $\lambda, \mu \in C_{\mathbb{Z}}$ by Lemma 8.1(ii). If G is not of type A_1, then $\mathrm{Ext}^1_{\mathfrak{g}}(L(\lambda), L(\mu)) = 0$ by Theorem 8.9 and the exceptional case, where $G = A_1$, is well known. □

As before there is a corollary:

COROLLARY 8.12. *Let G be a semisimple algebraic group and let V be a $\mathfrak{g}$-module with $p > \dim V$. Assume that $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$. Then V is semisimple.*

The next corollary uses a famous result of Serre on the semisimplicity of tensor products to extend our results a little further. This result will be crucial for showing the splitting of certain non-semisimple Lie algebras.

COROLLARY 8.13. *Let $\mathfrak{g}$ be a Lie algebra and V, W two semisimple $\mathfrak{g}$-modules with $\dim V + \dim W < p + 2$. Then $V \otimes W$ is semisimple.*
Furthermore, let $\mathfrak{g} = \mathrm{Lie}(G)$ for G a semisimple algebraic group with $p > 2$ and p very good. Then $\mathrm{H}^2(\mathfrak{g}, V \otimes W) = 0$ unless $\mathfrak{g}$ contains a factor $\mathfrak{sl}_2$ and $V \otimes W$

contains a composition factor of the $\mathfrak{sl}_2$*-module* $L(p-2)$*. Also* $\mathrm{H}^1(\mathfrak{g}, V\otimes W) = 0$*, unless one of* V *and* W *is isomorphic to* k *and we are in one of the exceptional case of Theorem 8.9.*

Proof. For the first statement, we begin with some reductions as in [Ser94]. If $W = 0$ or k there is nothing to prove. If W is at least 2-dimensional, then either $p = 2$ and V is trivial (so that the result holds), or both $\dim V$ and $\dim W < p$. We may assume that both V and W are simple. Further, we may replace $\mathfrak{g}$ by the restricted algebra generated by its image in $\mathfrak{gl}(V \oplus W)$. As $V \oplus W$ is a semisimple module, we may thus assume $\mathfrak{g}$ is p-reductive. Now $\mathfrak{g} \subseteq \mathfrak{gl}(V) \times \mathfrak{gl}(W) = \mathfrak{sl}(V) \times \mathfrak{sl}(W) \times \mathfrak{z}$, where $\mathfrak{z}$ is a torus, and where the projections of $\mathfrak{g}$ onto the first two factors are irreducible, hence semisimple by Theorem B. We thus may assume $\mathfrak{g} \subseteq \mathfrak{sl}(V) \times \mathfrak{sl}(W)$ is a semisimple restricted subalgebra.
By Theorem 2.3, either (i) $\mathfrak{g}$ has a factor W_1, the first Witt algebra and V is the $(p-1)$-dimensional irreducible module for W_1; or (ii) $\mathfrak{g}$ is $\mathrm{Lie}(G)$ for a direct product of simple algebraic groups, and V and W are (the differentials of) p-restricted modules for G. In case (i), as $p > 2$, we would have $W \cong k \oplus k$ for W_1 and the result holds. So we may assume that (ii) holds. Now [Ser94, Prop. 7] implies that $V \otimes W$ is the direct sum of simple modules with restricted high weights λ satisfying $\lambda \in C_{\mathbb{Z}}$. Since each of these composition factors is simple also for $\mathfrak{g}$, $V \otimes W$ is semisimple with those same composition factors.
For the remaining statements, let $\mathfrak{h}$ be the image of $\mathfrak{g}$ in $\mathfrak{gl}(V \oplus W)$, so that $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{s}$ with $\mathfrak{s}$ acting trivially. Let h be the coxeter number of $\mathfrak{h}$. Now if $W = k$, say, then since p is very good for $\mathfrak{g}$ we can have $p = \dim V$ by Proposition 8.1 only for $p > h$, so otherwise $\dim V < p$. And if $\dim W > 1$ then $\dim V < p$ also. Now $\dim V < p$ also implies by Proposition 8.1 that $p > h$. Also a summand $L(\lambda)$ of $V \otimes W$ has $\lambda \in C_{\mathbb{Z}}$. Now Theorem 8.9 implies that $\mathrm{H}^1(\mathfrak{g}, V \otimes W) = \mathrm{H}^2(\mathfrak{g}, V \otimes W) = 0$, unless we are in the exceptional cases described. However, if $\mathfrak{g} = \mathfrak{sl}_3$ then the module $L(p-3, 0)$ or its dual has dimension $(p-1)(p-2)(p-3)/2 > ((p+1)/2)^2$ hence it cannot appear as a composition factor of $V \otimes W$. ☐

Remark 8.14. If $\mathfrak{g} = W_1$ the conclusion of the second part is false, since $H^1(\mathfrak{g}, V) \neq 0$ when V is the irreducible $(p-1)$-dimensional module for $\mathfrak{g}$.

Proof of Theorem D:. We must just give references for the statements made. For (a), see Proposition 8.6; for (b), see Proposition 8.11; for (c), see Theorem 8.9; for (d), see Corollary 8.13. This completes the proof of Theorem D. ☐

9 Decomposability: the existence of Levi factors

Let $\mathfrak{h}$ be a restricted subalgebra of $\mathfrak{gl}(V)$ with $p > \dim V$. In this section we show, in Theorem 9.2, a strong version of the Borel–Tits Theorem in this context.

Let G be connected reductive. Recall, say from [ABS90] that if $\mathfrak{p} = \mathfrak{l} + \mathfrak{q}$ is a parabolic subalgebra of $\mathfrak{g} = \operatorname{Lie} G$ then $\mathfrak{q}$ has a central filtration such that successive quotients have the structure of $\mathfrak{l}$-modules. We record a specific case:

LEMMA 9.1. *In case $G = \mathrm{GL}_n$, a parabolic subalgebra $\mathfrak{p} = \mathfrak{l} + \mathfrak{q}$ has the property that $\mathfrak{l}$ is a direct product $\mathfrak{gl}(V_1) \times \mathfrak{gl}(V_2) \times \cdots \times \mathfrak{gl}(V_r)$ and $\mathfrak{q}$ has a central filtration with successive factors being modules of the form $V_i \otimes V_j^*$, each factor occurring exactly once.*

THEOREM 9.2. *Let $\mathfrak{h}$ be a restricted Lie subalgebra of $\mathfrak{gl}(V)$ with $\dim V < p$, and let $\mathfrak{r} = \mathrm{Rad}_p(\mathfrak{h})$ $(= \mathrm{Rad}_V(\mathfrak{h}))$.*
Then there is a parabolic subalgebra $\mathfrak{p} = \mathfrak{l} + \mathfrak{q}$, with $\mathfrak{r} \leq \mathfrak{q}$ and containing a complement $\mathfrak{s}$ to $\mathfrak{r}$ in $\mathfrak{h}$, with $\mathfrak{s} \leq \mathfrak{l}$ and $\mathfrak{h} = \mathfrak{s} + \mathfrak{r}$ as a semidirect product. Furthermore, $\mathfrak{s}$ acts completely reducibly on V and is the direct sum of a torus and a semisimple ideal.

Proof. As in the proof of Lemma 7.9 we take a minimal parabolic subgroup $P = LQ$ so that its Lie algebra $\mathfrak{p} = \mathfrak{l} + \mathfrak{q}$ contains $\mathfrak{h}$ and so that the projection $\mathfrak{h}_\mathfrak{l} := \pi(\mathfrak{h})$ of $\mathfrak{h}$ to the Levi subalgebra $\mathfrak{l}$ is strongly p-reductive and we may write $\mathfrak{h}_\mathfrak{l} = \mathfrak{h}_s \oplus \mathfrak{z}$ where $\mathfrak{h}_s$ is semisimple and $\mathfrak{z} = Z(\mathfrak{h}_\mathfrak{l})$. We also have $\mathfrak{r} \leq \mathfrak{q}$, since $\mathfrak{h}_\mathfrak{l}$ is p-reductive.
Now by Theorem 2.3, either $\mathfrak{h}_s = W_1$, $\mathfrak{h} = \mathfrak{h}_\mathfrak{l}$, $\mathfrak{p} = \mathfrak{l} = \mathfrak{gl}(V)$ and we are done; or $\mathfrak{h}_\mathfrak{l}$ is isomorphic to a direct product of classical Lie algebras $\mathfrak{s}_i$ and $\mathfrak{z}$.
We first lift $\mathfrak{z}$ to $\mathfrak{h}$. Let $\pi' : \mathfrak{h} \to \mathfrak{z}$ be the composition of π with the projection onto $\mathfrak{z}$. By [SF88, Lemma 2.4.4(2)], there is a torus $\mathfrak{z}' \leq Z(\mathfrak{l}) + \mathfrak{q}$ so that $\mathfrak{h} = \mathfrak{z}' + \ker(\pi')$. Now since $\mathfrak{z}'$ is a torus, it is linearly reductive, we may replace $\mathfrak{h}$ by a conjugate by Q so that $\mathfrak{z}' \subseteq Z(\mathfrak{l})$. Let us rewrite $\mathfrak{z} = \mathfrak{z}'$ and identify $\mathfrak{z}$ with its image in $\mathfrak{l}$ under π.
Next we construct a complement to $\mathfrak{r}$ in $\mathfrak{h}$. Let $\pi'' : \mathfrak{h} \to \mathfrak{h}_s$ be the composition of π with the projection onto $\mathfrak{h}_s$ and let $\mathfrak{h}' \subseteq \mathfrak{h}$ be a vector space complement to $\ker(\pi'')$. Then $\mathfrak{r} + \mathfrak{h}' \leq \mathfrak{h}$ is a subalgebra, and we have an exact sequence

$$0 \to \mathfrak{r} \to \mathfrak{r} + \mathfrak{h}' \xrightarrow{\pi''} \mathfrak{h}_s \to 0.$$

We show this sequence is split. By Lemma 9.1, the nilpotent radical $\mathfrak{q}$ of $\mathfrak{l}$ has a filtration $\mathfrak{q} = \mathfrak{q}_1 \supseteq \mathfrak{q}_2 \supseteq \cdots \supseteq \mathfrak{q}_m = 0$ with each $\mathfrak{q}_i/\mathfrak{q}_{i+1}$ having the structure of an $\mathfrak{l}$-module $M_i \otimes N_i$ with M_i and N_i irreducible modules for the projections of $\mathfrak{h}_\mathfrak{l}$ to distinct factors of the Levi. Since $\dim M_i + \dim N_i < p$, we have by Corollary 8.13 that $M_i \otimes N_i$ is a direct sum of irreducible modules for $\mathfrak{h}_s$ with $\mathrm{H}^2(\mathfrak{h}_s, M_i \otimes N_i) = 0$. By intersecting with $\mathfrak{r}$, we get a filtration $\mathfrak{r} = \mathfrak{r}_1 \supseteq \mathfrak{r}_2 \supseteq \cdots \supseteq \mathfrak{r}_m = 0$ by $\mathfrak{h}_s$-modules so that each $\mathfrak{r}_i/\mathfrak{r}_{i+1}$ is a submodule of $M_i \otimes N_i$, hence also a semisimple module with $\mathrm{H}^2(\mathfrak{h}_s, \mathfrak{r}_i/\mathfrak{r}_{i+1}) = 0$. By an obvious induction on the length m of the filtration $\{\mathfrak{r}_i\}$ we now see that the sequence

$$0 \to \mathfrak{r} \to \mathfrak{r} + \mathfrak{h}' \to \mathfrak{h}_s \to 0$$

is split. Thus we may set $\mathfrak{h}'_s$ a complement to $\mathfrak{r}$ in $\mathfrak{h}' + \mathfrak{r}$.

We would like to set $\mathfrak{s} = \mathfrak{h}'_s + \mathfrak{z}$, however this vector space may not be a subalgebra of $\mathfrak{g}$. Write $\mathfrak{q} = \mathfrak{c}_{\mathfrak{q}}(\mathfrak{z}) + [\mathfrak{q}, \mathfrak{z}]$. (This can be done, for instance by [SF88, Lemma 2.4.4(1)].) Any element h of $\mathfrak{h}'_s$ can be written as $h_1 + q_1 + q_2$ for $h_1 \in \mathfrak{l}$, q_1 in $\mathfrak{c}_{\mathfrak{q}}(\mathfrak{z})$ and $q_2 \in [\mathfrak{q}, \mathfrak{z}]$. As $\mathfrak{h}$ is stable under $\operatorname{ad}\mathfrak{z}$, with $\mathfrak{z}$ centralising h_1 and q_1, we conclude that $q_2 \in \mathfrak{h}$. Thus we have the element $h' = h_1 + q_1 \in \mathfrak{h}$. Thus we may form the subspace $\mathfrak{h}''_s \leq \mathfrak{h}$ with $\mathfrak{h}''_s \leq \mathfrak{l} + \mathfrak{c}_{\mathfrak{q}}(\mathfrak{z})$.
Using that $\mathfrak{h}'_s \leq \mathfrak{h}$ is a subalgebra, that $\mathfrak{c}_{\mathfrak{q}}(\mathfrak{z})$ is $\mathfrak{l} = \mathfrak{c}_{\mathfrak{gl}(V)}(\mathfrak{z})$-invariant and that $[\mathfrak{q}, \mathfrak{z}]$ is an ideal in $\mathfrak{q}$, one checks that $\mathfrak{h}''_s$ is indeed a subalgebra,[4] with $\mathfrak{h}''_s$ also a complement to $\mathfrak{r}$ in $\mathfrak{h}'_{\mathfrak{s}} + \mathfrak{r}$. Now we have guaranteed that $\mathfrak{s} = \mathfrak{h}''_s + \mathfrak{z}$ is a subalgebra of $\mathfrak{h}$, a complement to $\mathfrak{r}$ in $\mathfrak{h}$.
Now, by Corollary 8.12, $\mathfrak{h}''_s$ acts completely reducibly. Also, since $\mathfrak{z}$ is a torus, $\mathfrak{z}$ is linearly reductive on restricted representations, hence also acts completely reducibly. Thus $\mathfrak{s}$ is completely reducible on V. In particular, we may replace $\mathfrak{l}$ with a Levi subalgebra of $\mathfrak{p}$ that contains $\mathfrak{s}$, which finishes the proof. $\square$

10 Proof of Theorem B(i)

Proof. We first prove the statement in the case that $G = \mathrm{GL}(V)$, so we assume $p > \dim V + 1$. By assumption, $\mathfrak{h}$ is a restricted subalgebra of $\mathfrak{g}$.
Let $\mathfrak{n} = \mathfrak{n}_{\mathfrak{g}}(\mathfrak{h})$. By Theorem 9.2 we may decompose both $\mathfrak{n}$ and $\mathfrak{h}$. Let $\mathfrak{n} = \mathfrak{n}_{\mathfrak{l}} + \mathfrak{n}_{\mathfrak{q}} \leq \mathfrak{p} = \mathfrak{l} + \mathfrak{q}$ with $\mathfrak{n}_{\mathfrak{l}} \leq \mathfrak{l}$ and $\mathfrak{n}_{\mathfrak{q}} \leq \mathfrak{q}$, with $\mathfrak{n}_{\mathfrak{l}} = \mathfrak{n}_s + \mathfrak{z}$, $\mathfrak{z}$ a torus and $\mathfrak{n}_s$ is by Theorem 2.3 isomorphic to a direct product of classical Lie algebras acting completely reducibly on V; also set $\mathfrak{h}_{\mathfrak{q}} = \mathfrak{h} \cap \mathfrak{q}$ and $\mathfrak{h}_{\mathfrak{l}} = \pi(\mathfrak{h})$ the projection to $\mathfrak{l}$. Since $\mathfrak{n}$ is generated by nilpotent elements we have $\mathfrak{z} = 0$ and $\mathfrak{h}_{\mathfrak{l}} = \mathfrak{h}_s$. Since the complement to $\mathfrak{h}_{\mathfrak{q}}$ in $\mathfrak{h}$ obtained by Theorem 9.2 is completely reducible on V and hence conjugate to a subalgebra of $\mathfrak{l}$, we may assume that $\mathfrak{h} = \mathfrak{h}_{\mathfrak{q}} + \mathfrak{h}_{\mathfrak{l}}$ is this splitting. Furthermore, $\mathfrak{h}_{\mathfrak{l}} \leq \mathfrak{n}_{\mathfrak{l}}$ is an ideal of a direct product of simple subalgebras, hence is a direct product of some subset of those simples.
Since V has dimension less than p, $V|_{\mathfrak{n}_{\mathfrak{l}}}$ is a restricted module for $\mathfrak{n}_{\mathfrak{l}}$. Hence there is a connected algebraic group $N_{\mathfrak{l}}$ with $\operatorname{Lie} N_{\mathfrak{l}} \cong \mathfrak{n}_{\mathfrak{l}}$, $N_{\mathfrak{l}} \leq \mathrm{GL}(V)$ and $V|_{\mathrm{Lie}(N_{\mathfrak{l}})} \cong V|_{\mathfrak{n}_{\mathfrak{l}}}$. Hence, replacing $N_{\mathfrak{l}}$ by a conjugate if necessary, we have $\mathrm{Lie}(N_{\mathfrak{l}}) = \mathfrak{n}_{\mathfrak{l}}$. Moreover if L is a Levi subgroup of $\mathrm{GL}(V)$ chosen so that $\mathrm{Lie}(L) = \mathfrak{l}$ then we may produce $N_{\mathfrak{l}} \leq L$. Clearly $N_{\mathfrak{l}}$ normalises any direct factor of $\mathfrak{n}_{\mathfrak{l}}$, in particular, $\mathfrak{h}_{\mathfrak{l}}$.
Now, since the $\mathfrak{l}$-composition factors of $\mathfrak{q}$ are all of the form $W_1 \otimes W_2$ for $\dim W_1 + \dim W_2 < p$ and W_1, W_2 irreducible for $\mathfrak{n}_{\mathfrak{s}}$, [Ser94, Prop. 7] implies that $\mathfrak{q}$ is a restricted semisimple module for $N_{\mathfrak{l}}$ and $\mathfrak{n}_{\mathfrak{l}}$. Since $\mathfrak{n}_{\mathfrak{l}}$ normalises $\mathfrak{h}_{\mathfrak{q}} = \mathfrak{h} \cap \mathfrak{q}$, this space also appears as an $N_{\mathfrak{l}}$-submodule in $\mathfrak{q}$, hence $N_{\mathfrak{l}}$ normalises $\mathfrak{h}_{\mathfrak{q}}$.

[4]The calculation is as follows: if $h_1 + q_1 + q_2$ and $h'_1 + q'_1 + q'_2$ are two elements of $\mathfrak{h}'_s$ then

$$[h_1 + q_1 + q_2, h'_1 + q'_1 + q'_2] = \underbrace{[h_1, h_2]}_{\in \mathfrak{h}_{\mathfrak{l}}} + \underbrace{[h_1, q'_1] + [q_1, h'_1] + [q_1, q'_1]}_{\in \mathfrak{c}_{\mathfrak{q}}(\mathfrak{z})} + x,$$

where $x \in [\mathfrak{q}, \mathfrak{z}]$ by the Jacobi identity. But projecting to $\mathfrak{h}''_s$ one simply deletes q_2, q'_2 and x to get the analagous calculation.

It remains to construct a unipotent algebraic group $N_{\mathfrak{q}}$ such that $\operatorname{Lie} N_{\mathfrak{q}} = \mathfrak{n}_{\mathfrak{q}}$ with $N_{\mathfrak{q}}$ normalising $\mathfrak{h}$. For this we use Corollary 4.4. Let $N_{\mathfrak{q}} = \overline{\langle \exp x : x \in \mathfrak{n}_{\mathfrak{q}} \rangle}$. Then $N_{\mathfrak{q}}$ is a closed subgroup, which by Corollary 4.4 consists of elements normalising $\mathfrak{h}$. By Lemma 4.5, $\mathfrak{n}_{\mathfrak{q}} \leq \operatorname{Lie}(N_{\mathfrak{q}})$.
Let N be the smooth algebraic group given by $N = \langle N_{\mathfrak{l}}, N_{\mathfrak{q}} \rangle$. We have shown that N normalises $\mathfrak{h}$ and that $\mathfrak{n} \subseteq \operatorname{Lie} N$. Since also $\operatorname{Lie} N \subseteq \mathfrak{n}$ we are done for the case $G = \mathrm{GL}(V)$.
To prove the remaining part, we appeal to Proposition E again.
Let G be a simple algebraic group with minimal dimensional representation V. Then since $p > \dim V$, $(\mathrm{GL}(V), G)$ is a reductive pair. Indeed, the assumption on p guarantees that the trace form associated to V is non-zero, see [Gar09, Fact 4.4]. This implies the reductive pair property (cf. the proof of [Gar09, Prop. 8.1]). The theorem now follows by invoking Proposition E. ☐

11 Examples

In this section we mainly collect, in a number of statements, examples which demonstrate the tightness of some of our bounds. First let us just point out that there are some rather general situations in which smooth normalisers can be found.

Example 11.1 ([MT09, Theorem B]). Suppose G is a quasi-split reductive group over a field k of very good characteristic. Then the normaliser $N = N_G(C)$ of the centraliser $C = C_G(e)$ of a regular nilpotent element e of $\mathfrak{g} = \operatorname{Lie}(G)$ is smooth.

Example 11.2 ([HS16, Proof of Lem. 3.1]). Suppose G is reductive over an algebraically closed field k of very good characteristic and e is a nilpotent element of $\mathfrak{g} = \operatorname{Lie}(G)$, then the normaliser $N_G(\langle e \rangle)$ of the 1-space $\langle e \rangle$ of $\mathfrak{g}$ is smooth.

We will first give the promised example discussed after the statement of Theorem A. For this, we will need a lemma.

Lemma 11.3. *Let $B = TU$ be a Borel subgroup of a reductive algebraic group G containing a maximal torus T with unipotent radical U. Suppose $N_B(\mathfrak{h})$ is smooth and $s \in \mathfrak{t} = \operatorname{Lie}(T)$ an element normalising a subspace $\mathfrak{h}$ of $\mathfrak{u} = \operatorname{Lie}(U)$. Then $\langle s \rangle = \operatorname{Lie}(\chi(\mathbb{G}_m))$ for a cocharacter $\chi : \mathbb{G}_m \to N_B(\mathfrak{h})$, such that $\chi(\mathbb{G}_m)$ is conjugate by an element of $C_U(s)$ to a cocharacter with image in T.*

Proof. Since $N_B(\mathfrak{h})$ is smooth, we may, by [Hum67, Thm. 13.3], write any maximal torus $\mathfrak{s}$ of $\mathfrak{n}_{\mathfrak{b}}(\mathfrak{h})$ as $\operatorname{Lie}(S)$ for S a maximal torus of $N_B(\mathfrak{h})$. By [Die52, Prop. 2], for any semisimple element $s \in \mathfrak{s}$ we may write $\langle s \rangle = \operatorname{Lie}(S_1)$ for $S_1 \subseteq S$. Defining an appropriate isomorphism $\mathbb{G}_m \to S_1$, we may even write $s = \frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=1} \chi(t)$ for χ a cocharacter of $N_B(\mathfrak{h})$.
As the maximal tori of B are conjugate by elements of U, we have that S is conjugate to its projection to T, say via $u \in U$; in particular, $u\chi(t)u^{-1} \in$

T. Since projection to T is B-equivariant, we have on differentiating, that $\frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=1}(u\chi(t)u^{-1}) = s$, so that $usu^{-1} = s$, i.e. that $u \in C_U(s)$. $\square$

Example 11.4. Let $n \geq 4$. This example depends on three fixed parameters $\lambda_1, \lambda_2, \lambda_3$ together with variables $\{a_i\}_{1\leq i\leq n}$, $\{b_i\}_{1\leq i\leq n-1}$, c, d, and e, each taking values in $k = \bar{\mathbb{F}}_p$.
Let us define the following matrices:

$$A := \begin{pmatrix} 0 & a_1 & a_2 & * & * & * & \dots & * & * & * \\ & 0 & a_1 & \beta_2 & * & * & \dots & * & * & * \\ & & 0 & a_2 & \beta_3 & * & \dots & * & * & * \\ & & & 0 & a_3 & \beta_4 & \ddots & \vdots & \vdots & \vdots \\ & & & & 0 & a_4 & \ddots & * & * & * \\ & & & & & 0 & \ddots & \beta_{n-2} & c & e + \lambda_1\beta_{n-2} \\ & & & & & & \ddots & a_{n-2} & \beta_{n-1} & (1+\lambda_1)a_{n-2} + d \\ & & & & & & & 0 & a_{n-1} & b_{n-1} \\ & & & & & & & & 0 & a_n \\ & & & & & & & & & 0 \end{pmatrix},$$

with $\beta_i = a_{i+1} + b_{i-1}$ for $i = 2, \dots, n-1$,

$$B := \begin{pmatrix} 0 & a_1 & b_1 & * & \dots & * \\ & 0 & a_2 & b_2 & \ddots & \vdots \\ & & 0 & a_3 & \ddots & * \\ & & & 0 & \ddots & b_{n-1} \\ & & & & \ddots & a_n \\ & & & & & 0 \end{pmatrix},$$

$$C := \begin{pmatrix} 0 & a_{n-3} & a_{n-1} + b_{n-3} & c & e + \lambda_2(a_{n-1} + b_{n-3}) \\ & 0 & a_{n-2} & a_n + b_{n-2} & (1+\lambda_2)a_{n-2} + d \\ & & 0 & a_{n-1} & b_{n-1} \\ & & & 0 & a_n \\ & & & & 0 \end{pmatrix},$$

$$D := \begin{pmatrix} 0 & a_{n-2} & a_n + b_{n-2} & d + \lambda_3 a_{n-2} \\ & 0 & a_{n-1} & b_{n-1} \\ & & 0 & a_n \\ & & & 0 \end{pmatrix}$$

Then the reader may check that for each choice of λ_1, λ_2 and λ_3, the following set defines a subalgebra $\mathfrak{h}$ of the strictly upper triangular matrices:

$$\left\{ \begin{pmatrix} A & * & * & * \\ 0 & B & * & * \\ 0 & 0 & C & * \\ 0 & 0 & 0 & D \end{pmatrix} : a_i \in k, b_j \in k, c, d, e \in k \right\}.$$

Let F_i denote the ith Fibonacci number, so that $F_0 = F_1 = 1$ and $F_2 = 2$ and suppose that r is chosen so that $F_{r+1} = p$ is the prime characteristic of k, and let us suppose that $N_G(\mathfrak{h})$ is smooth. Since every entry of the superdiagonal is non-zero for some element in $\mathfrak{h}$, it is easy to check that $N_G(\mathfrak{h}) \subseteq B$. Thus $N_G(\mathfrak{h}) = N_B(\mathfrak{h})$ and we may employ Lemma 11.3.

Suppose $s = \operatorname{diag}(s_1, \ldots, s_{2n+12})$ is an arbitrary element of the diagonal torus $\mathfrak{t} = \operatorname{Lie}(T)$. Then one can calculate the dimension of $\mathfrak{n}_{\mathfrak{t}}(\mathfrak{h})$ by enumerating the linear conditions amongst the t_i necessary to normalise $\mathfrak{h}$. For example, setting all indeterminates in a general matrix of $\mathfrak{h}$ to be zero, except for $a_1 = 1$ gives a matrix M, which spans a 1-space $\langle M \rangle$ of $\mathfrak{h}$. One can see by inspection that s will normalise $\mathfrak{h}$ only if it normalises $\langle M \rangle$. However, calculating $[s, M]$, we see that to normalise $\langle M \rangle$ implies the following condition must hold:

$$s_1 - s_2 = s_2 - s_3 = s_{2n+4} - s_{2n+5}.$$

Repeating over other 1-spaces leads to a collection of relations which can be expressed by a system of linear equations $R\mathbf{s} = 0$ for some matrix R and the vector $\mathbf{s} = (s_1, \ldots, s_{2n+12})$. The kernel of R modulo p then determines the dimension of $\mathfrak{n}_{\mathfrak{t}}(\mathfrak{h})$. To determine the dimension of $N_T(\mathfrak{h})$, one searches for cocharacters $\chi(t) = \operatorname{diag}(t^{k_1}, t^{k_2}, \ldots, t^{k_{2n+12}})$ which normalise $\mathfrak{h}$ by conjugation. This leads to an identical set of relations on the entries of the vector $\mathbf{k} = (k_1, \ldots, k_{2n+12})$, so that the equation $R\mathbf{k} = 0$ must be solved *over the integers.* Then the dimension of $N_T(\mathfrak{h})$ is the nullity of R.

The nullities of R over $\mathbb{Z}$ and over $\mathbb{Z}/p$ are identical if and only if s can be lifted to a diagonal cocharacter $\chi(t)$ so that $d/dt|_{t=1}\chi(t) = s$. By explicit calculation of R in our particular case, one sees its elementary divisors are $0^4, 1^{2n+7}, F_{r+1}$. Thus since $p = F_{r+1}$ the nullity of R modulo p is bigger than over $\mathbb{Z}$. Thus there is a toral element s, which cannot be lifted to a diagonal cocharacter. In our case, $\mathfrak{h}$ has an obvious centraliser whose elements are:

$$\operatorname{diag}(\underbrace{s_1, \ldots, s_1}_{r+2}, \underbrace{s_2, \ldots, s_2}_{r+1}, \underbrace{s_3, \ldots, s_3}_{5}, \underbrace{s_4, \ldots, s_4}_{4})$$

which accounts also for the four-dimensional kernel over the integers.

One also checks that the subalgebra $\mathfrak{h}$ is normalised by the toral element

$$\begin{aligned} s :=& \operatorname{diag}(1, 2, 3, 5, 8, \ldots, F_r, F_{r+1}, F_{r+2}) \\ &\oplus \operatorname{diag}(F_2 + 4 = 6, F_3 + 4 = 7, \ldots, F_r + 4, F_{r+1} + 4, F_{r+2} + 4 = F_r + 4) \\ &\oplus \operatorname{diag}(F_{r-2} + 1, F_{r-1} + 1, F_r + 1, F_{r+1} + 1, F_{r+2} + 1) \\ &\oplus \operatorname{diag}(F_{r-1} + 2, F_r + 2, F_{r+1} + 2, F_{r+2} + 2), \end{aligned}$$

where for the direct sum $A \oplus B$ of two square matrices A and B we mean the block diagonal matrix having A and B on the diagonal. Note the congruence amongst the entries in t, $F_r = F_{r+2} \mod p$. Thus on each line, the last and pen-penultimate entries are the same modulo p. Furthermore, since this element does not centralise $\mathfrak{h}$, it can have no lift to a diagonal cocharacter.
By assumption, $N_G(\mathfrak{h}) = N_B(\mathfrak{h})$ is smooth. Thus $\langle s\rangle$ lifts to the image of a cocharacter χ' which, by Lemma 11.3 is conjugate by $C_U(s)$ to a diagonal cocharacer χ. Since by inspection, only five entries of s are the same, s is a regular toral element of $\mathfrak{n}_{\mathfrak{g}}(\mathfrak{h})$ and one checks

$$C_U(s) = \langle 1 + te_{r,r+2}, 1 + te_{2r+1,2r+3}, 1 + te_{2r+6,2r+8}, 1 + te_{2r+10,2r+12} : t \in k\rangle.$$

The action of the second listed element in turn normalises $\mathfrak{h}$ and the first, third and fourth simply change the values of $\lambda_1, \lambda_2, \lambda_3$. Thus if $g \in C_U(s)$ then one computes a new relation matrix R' computing the normaliser $\mathfrak{n}_{\mathfrak{t}}(\mathfrak{h}^g)$ which, by virtue of being independent of the values of λ_i, is identical to R. In particular, $\langle s\rangle$ still normalises $\mathfrak{h}$ but there is still no lift to a diagonal cocharacter. This contradicts the conclusion of Lemma 11.3, hence $N_G(\mathfrak{h})$ is not smooth.

The next example will show the necessity of the bound in Theorem 3.2. We first collect some miscellaneous auxiliary results in the following lemma. Recall that a subgroup H of a connected reductive group G is called *G-irreducible* if it is in no proper parabolic subgroup of G.

Lemma 11.5. *Suppose G is a connected reductive algebraic group and H is a (possibly disconnected) closed reductive subgroup of G.*
(i) We have $N_G(H)^\circ_{\mathrm{red}} = H^\circ C_G(H)^\circ_{\mathrm{red}}$.
(ii) If H is G-irreducible, then $C_G(H)^\circ_{\mathrm{red}} = \mathrm{Rad}(G)$, where $\mathrm{Rad}(G) = Z(G)^\circ_{\mathrm{red}}$.
(iii) Suppose $H \leq M \leq G$ is an intermediate reductive subgroup with $\mathrm{Rad}(G) \leq \mathrm{Rad}(M)$ and that H is G-irreducible. Write $Z(M)^\circ = \mathrm{Rad}(M) \times \mu_M$ for an infinitesimal subgroup scheme μ_M. Then either $\mu_M \leq Z(H)$ or $N_G(H)$ is non-smooth.

Proof. (i) and (ii) follow from [Mar03, Lemmas 6.2 and 6.8].
For (iii), clearly $\mu_M \leq Z(M) \leq N_G(H)$. If $N_G(H)$ is smooth, then by parts (i) and (ii) we have $\mu_M \leq Z(M)^\circ \leq H^\circ C_G(H)^\circ_{\mathrm{red}} = H^\circ \mathrm{Rad}(G)$. This forces $\mu_M \leq H^\circ$. ☐

Examples 11.6. Lemma 11.5 can be used to produce reductive subgroups H of G with non-smooth normalisers in bad characteristic. We use [Her13, Example 4.1], in which the first author constructs examples of non-smooth centralisers for each reductive group over a field of characteristic p for which p is not a very good prime for G. All the subgroups constructed in *loc. cit.* are maximal rank reductive subgroups M such that $C_G(M) = Z(M)$ is non-smooth, hence $\mu_M \neq 1$ in Lemma 11.5(iii) above. In many cases, we may take a further connected, reductive G-irreducible subgroup H of M such that p is pretty good for H. Thus its centre is in fact smooth, and being finite, cannot contain

μ_M. Thus by Lemma 11.5(iii) the normaliser $N_G(H)$ is non-smooth. Let us list some triples (G, p, M, H) which work for this process. By V_n we denote a natural module of dimension n for the classical group M; by $\tilde{M}_1$ we mean a subgroup of type M_1 corresponding to short roots.

G	p	M	H
G_2	3	A_2	$A_1 \hookrightarrow M;\ V_3\|_H = L(2)$
F_4	2	A_1^4	$A_1 \hookrightarrow M;\ x \mapsto (x, x^2, x^4, x^{16})$
F_4	3	$A_2\tilde{A}_2$	$(A_1, A_1) \hookrightarrow M; (V_3, V_3)\|_H = (L(2), L(2))$
E_8	5	A_4^2	$A_1^2 \hookrightarrow M;\ (V_5, V_5)\|_H = (L(4), L(4))$
SL_p	$p > 2$	SL_p	$A_1 \hookrightarrow M;\ V_p\|_H = L(p-1)$.

Remark 11.7. A complete list of conjugacy classes of simple G-irreducible subgroups of exceptional groups has been compiled by A. Thomas, see [Tho15] for the cases of rank at least 2 and [Tho16] for the rank 1 case. For the G_2 example one may consult [Ste10, Theorem 1, Corollary 3].

The next example shows the promised tightness of Theorem B(i) as stated in Remark 1.2(a).

Lemma 11.8. *Let $G = \mathrm{GL}(V)$ with $\dim V \geq p-1 \geq 3$ and take any subspace $W \leq V$ with $\dim W = p-1$. Then if $W_1 \leq \mathfrak{gl}(W)$ is the first Witt algebra in its $p-1$-dimensional representation we have $N_G(W_1)$ is not smooth.*

Proof. Since W_1 is irreducible on W, the normaliser $\mathfrak{n}_{\mathfrak{gl}(V)}(W_1) = \mathfrak{n}_{\mathfrak{sl}(W)}(W_1) \oplus \mathfrak{z} \oplus \mathfrak{gl}(U)$ for $V = W \oplus U$ and $\mathfrak{z}$ the centre of $\mathfrak{gl}(W)$. Moreover as W_1 is irreducible on W, so is $\mathfrak{n} = \mathfrak{n}_{\mathfrak{sl}(W)}(W_1)$. By Theorem C, $\mathfrak{n}$ is semisimple, hence, as W_1 is simple, it must be a direct factor of $\mathfrak{n}$, say $\mathfrak{n} = W_1 \oplus \mathfrak{h}$. But now the action of $\operatorname{ad}\mathfrak{h}$ on W is a W_1-module homomorphism, hence is a scalar by Schur's lemma. Thus $\mathfrak{h} \leq \mathfrak{z}(\mathfrak{sl}(W)) = 0$. It follows that $\mathfrak{n} = W_1$.
Now $N_G(W_1)$ sends W to another W_1-invariant subspace of the same dimension, hence $N_G(W_1) \leq \mathrm{GL}(W) \times \mathrm{GL}(U)$. Since W_1 is self-normalising, if $N_G(W_1)$ were smooth we would have $\operatorname{Lie} N_G(W_1) = \mathfrak{n}_{\mathfrak{g}}(W_1) = W_1 \oplus \mathfrak{gl}(U)$. This shows that W_1 is algebraic, a contradiction. □

We now justify the remark after Theorem B that the bound in Theorem B(i) is tight for $G = \mathrm{Sp}_{2n}$.

Lemma 11.9. *The p-dimensional Witt algebra W_1 is a maximal subalgebra of $\mathfrak{sp}_{p-1}$. Furthermore, its normaliser in any Sp_{p-1}-Levi of Sp_{2n} with $2n \geq p-1$ is non-smooth.*

Proof. Since W_1 stabilises the element

$$X \wedge X^{p-1} + \frac{1}{2}X^2 \wedge X^{p-2} + \frac{1}{3}X^3 \wedge X^{p-3} + \cdots + \frac{2}{p-1}X^{(p-1)/2} \wedge X^{(p+1)/2} \in \bigwedge^2 V$$

we find that W_1 is contained in $\mathfrak{sp}_{p-1}$, acting irreducibly on the $p-1$-dimensional module. Exponentiating a set of nilpotent generators of the

Witt algebra as in the proof of Theorem B(ii) gives an irreducible subgroup $W \leq \mathrm{Sp}_{p-1}$. We claim that we must have equality. From this claim it follows that W_1 is in no proper classical algebraic subalgebra of $\mathfrak{sp}_{p-1}$, hence, by Theorem 2.3, is maximal.
To prove the claim, suppose W is a proper subgroup of $G = \mathrm{Sp}_{p-1}$. Since W is irreducible on the $p-1$-dimensional module, W is it no parabolic of G. Thus it is in a connected reductive maximal subgroup M. We must have that M is simple, or else W_1 would be in a parabolic of G. Now since the lowest dimensional non-trivial representation of W_1 is $p-1$, it follows that M can have no lower-dimensional non-trivial representation. Since $p > 2$, Sp_{p-1} has no simple maximal rank subgroup. All classical groups of rank lower than $\frac{p-1}{2}$ have natural modules of smaller dimension than $p-1$, so M is of exceptional type. The lowest dimensional representations of the exceptional types are 6 ($p = 2$), 7, 25 ($p = 3$), 26, 27, 56 and 248. The only time one of these is $p-1$ is when $p = 57$ and $M = E_7$. But if $p = 57$ then $p > 2h - 2$ for E_7, then by Theorem B(ii) all maximal semisimple subalgebras are algebraic and so W_1 is not a subalgebra of E_7. This proves the claim, hence gives the first part of the lemma.
For the second, with $2n > p-1$, we have $W_1 \leq \mathfrak{sp}_{p-1} \oplus \mathfrak{sp}_{2n-p+1}$ with W_1 sitting in the first factor. Then its normaliser is evidently $W_1 \oplus \mathfrak{sp}_{2n-p+1}$, however this is not algebraic for $p > 3$, hence the normaliser $N_G(W_1)$ cannot be smooth. Thus we have shown that normalisers of all subalgebras of $\mathfrak{sp}_{2n}$ are smooth only if $p > h + 1$. □

If $n \geq p$ there is a more straightforward example of a (non-restricted) subalgebra of $\mathfrak{gl}_n$ whose normaliser in GL_n is not smooth.

Example 11.10. Let $\mathfrak{g} = \mathfrak{gl}_n$, take J_p a Jordan block of size p and take the abelian one-dimensional Lie algebra $\mathfrak{h} = k(I_p + J_p)$ where I_p is an identity block of size p. Then one can show with elementary matrix calculations that the normaliser of $N_G(\mathfrak{h})$ is non-smooth.

The next example shows that even the normalisers of smooth groups are not smooth, even in $\mathrm{GL}(V)$, and even when p is arbitrarily large.

Lemma 11.11. *Let $G = \mathrm{GL}(V)$ with $\dim V \geq 3$ and let W be a 3-dimensional subspace. Let $U \leq \mathrm{GL}(W)$ be defined as the smooth subgroup whose k-points are*

$$U(k) = \left\{ \begin{bmatrix} 1 & 0 & t \\ 0 & 1 & t^p \\ 0 & 0 & 1 \end{bmatrix} : t \in k \right\}.$$

Write $V = W \oplus W'$ for some complement W' to W and set $H = U \oplus \mathrm{GL}(W') \leq \mathrm{GL}(V)$. Then $N_G(H)$ is non-smooth.

Proof. From the reductivity of $\mathrm{GL}(W')$ it follows that $N_G(H) = N_{\mathrm{GL}(W)}(U) \oplus \mathrm{GL}(W')$ so it suffices to show that $N_{\mathrm{GL}(W)}(U)$ is non-smooth. This is a routine calculation. For example, if x is an element of a k-algebra A, with $x^p = 0$ then one checks that the matrix

$$\begin{bmatrix} 1 & x & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \in N_{\mathrm{GL}(W)}(U)(A).$$

Now, the normaliser of U of course normalises $\mathrm{Lie}(U)$. Since

$$\mathrm{Lie}(U) = k \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

the normaliser of $\mathrm{Lie}(U)$ is the product of the centraliser of a certain (nilpotent) element and the image of a cocharacter associated with that element. In particular, the normaliser of $\mathrm{Lie}(U)$ is contained in the upper triangular Borel subgroup.
Write V for the unipotent radical of that Borel subgroup, so V is 3-dimensional; a typical element has the form

$$\begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix}.$$

In fact, the condition $a^p = 0$ defines the scheme-theoretic normaliser in V of U, and the condition $a = 0$ defines the corresponding smooth subgoup of V whose k-points form the group-theoretic normaliser of $U(k)$ in $V(k)$. The lemma follows. $\square$

Now we show that normalisers of height two or more subgroup schemes are not smooth.

EXAMPLE 11.12. Let G be any connected reductive algebraic group over an algebraically closed field k of characteristic $p > 2$ and set $F : G \to G$ to be the Frobenius endomorphism. Let $B = TU$ be a Borel subgroup of G with T an F-stable maximal torus, and let U the non-trivial unipotent radical. Let T_r be the kernel in T of F^r and U_1 the kernel in U of F. Finally set $H = T_r \ltimes U_1$. Then $N_G(H) = T \ltimes U_1$, hence is not smooth.

The next example shows that if $p = \dim V$, then the normaliser of a smooth connected solvable non-diagonalisable algebraic subgroup of $\mathrm{GL}(V)$ can even be irreducible on V, thus *a fortiori* it is not smooth. This also gives an example for when $p = 2$ and $\dim V = 2$ that the normalisers in $\mathrm{SL}(V)$ and $\mathrm{GL}(V)$ of subalgebras of the respective Lie algebras are not smooth.

EXAMPLE 11.13. By [Ten87, Lemma 3] the Lie algebra $W_1 + O_1$ formed as the semidirect product of W_1 and O_1, where O_1 acts on itself by multiplication, is a maximal subalgebra of $\mathfrak{sl}_p = \mathfrak{sl}(k[X]/X^p)$. We imitate the embedding of

O_1 in $\mathfrak{gl}_p$ by a solvable subgroup of GL_p. Define the height $\mathrm{ht}(\alpha)$ of a root α to be the sum of the coefficients of the simple roots. Let U be the subgroup $\langle\prod_{\alpha\in R^-;\mathrm{ht}(\alpha)=i} x_\alpha\rangle_{1\le i\le p-1}$. By construction U is connected and unipotent and one can show that $\dim U = p-1$ and that $\mathrm{Lie}\, H = O_1$, where H is the smooth solvable subgroup $Z(\mathrm{GL}_p)U$. Now it can be shown that there is a subgroup scheme W corresponding to W_1 in GL_p which normalises H and for which $W \ltimes H$ is irreducible. It immediately follows that $N_G(H)$ cannot be smooth.

Finally we show that if $p \le 2n-1$ the normalisers in GL_n and SL_n of subspaces of their Lie algebras are not all smooth, even when these normalisers are generated by nilpotent elements, showing that the bound in Theorem B(ii) cannot be improved for general subspaces.

Lemma 11.14. *If $p < 2n-1$, normalisers of subspaces of $\mathfrak{gl}_n$ (or $\mathfrak{sl}_n$) are not necessarily smooth.*

Proof. Let $p = 2n-3$ and let $\mathfrak{h} = \mathfrak{sl}_2 = \mathrm{Lie}\, H$ with $H = \mathrm{SL}_2$ over a field k of characteristic p. Then the action of H on the simple module $L((p+1)/2)$ gives an (irreducible) embedding $H \to \mathrm{GL}_n$. Restricting the adjoint representation of $\mathfrak{gl}_n$ on itself to H gives a module

$$L((p+1)/2) \otimes L((p+1)/2)^* \cong T(p+1) \oplus M,$$

where M is a direct sum of irreducibles for H (and $\mathfrak{h}$) and $T(p+1)$ is a tilting module, uniserial with successive composition factors $L(p-3)|L(p+1)|L(p-3)$. Now for the algebraic group $H = SL_2$ we have $L(p+1) \cong L(1) \otimes L(1)^{[1]}$ by Steinberg's tensor product formula. Restricting to $\mathfrak{h}$, $L(p+1)$ is isomorphic to $L(1) \oplus L(1)$. Now it is easy to show the restriction map $\mathrm{Ext}^1_G(L(p+1), L(p-3)) \to \mathrm{Ext}^1_{\mathfrak{g}}(L(1), L(p-3)) \oplus \mathrm{Ext}^1_{\mathfrak{g}}(L(1), L(p-3))$ is injective. Hence $T(p+1)|_{\mathfrak{g}}$ contains a submodule M isomorphic to $L(1)/L(p-3)$.
Now, the Lie theoretic normaliser of M contains $\mathfrak{h}$ but the scheme-theoretic stabiliser does not contain H. It follows that the normaliser of this subspace is not smooth.
Indeed, as $\mathfrak{h}$ acts irreducibly on the n-dimensional natural representation for $\mathfrak{gl}_n$, it is in no parabolic of $\mathfrak{gl}_n$ (or $\mathfrak{sl}_n$). However, the set of k-points $N_H(M)(k) = N_{\mathrm{GL}_n}(M)(k) \cap H$ is in a parabolic of H, hence in a parabolic of GL_n. $\square$

References

[ABS90] H. Azad, M. Barry, and G. Seitz, *On the structure of parabolic subgroups*, Comm. Algebra 18 (1990), no. 2, 551–562. MR MR1047327 (91d:20048)

[BHMR11] M. Bate, S. Herpel, B. Martin, and G. Röhrle, *G-complete reducibility and semisimple modules*, Bull. Lond. Math. Soc. 43 (2011), no. 6, 1069–1078. MR 2861529 (2012k:20089)

[BNP02] C. P. Bendel, D. K. Nakano, and C. Pillen, *Extensions for finite Chevalley groups. II*, Trans. Amer. Math. Soc. 354 (2002), no. 11, 4421–4454 (electronic). MR 1926882 (2003k:20063)

[BNP04] ——, *Extensions for Frobenius kernels*, J. Algebra 272 (2004), no. 2, 476–511. MR 2028069 (2004m:20089)

[BNP07] ——, *Second cohomology groups for Frobenius kernels and related structures*, Adv. Math. 209 (2007), no. 1, 162–197. MR 2294220 (2008c:20085)

[Bou05] N. Bourbaki, *Lie groups and Lie algebras. Chapters 7–9*, Elements of Mathematics (Berlin), Springer-Verlag, Berlin, 2005, Translated from the 1975 and 1982 French originals by Andrew Pressley. MR 2109105 (2005h:17001)

[BT65] A. Borel and J. Tits, *Groupes réductifs*, Inst. Hautes Études Sci. Publ. Math. (1965), no. 27, 55–150. MR 0207712 (34 #7527)

[Del13] P. Deligne, *Semi-simplicité de produits tensoriels en caractéristique p*, Invent. Math. (2013), 1–25 (French).

[DG70] M. Demazure and P. Gabriel, *Groupes algébriques. Tome I: Géométrie algébrique, généralités, groupes commutatifs*, Masson & Cie, Éditeur, Paris, 1970, Avec un appendice ıt Corps de classes local par M. Hazewinkel.

[DGd70] M. Demazure and A. Grothendieck (dirig.), *Schémas en groupes, Séminaire de Géométrie Algébrique du Bois Marie 1962/64 (SGA 3)*, (Lecture Notes in Math., 151–153), Springer-Verlag, Berlin, 1970.

[Die52] J. Dieudonné, *Sur les groupes de Lie algébriques sur un corps de caractéristique* $p > 0$, Rend. Circ. Mat. Palermo (2) 1 (1952), 380–402 (1953). MR 0055355 (14,1062a)

[Dzh92] A. S. Dzhumadil′daev, *Cohomology and nonsplit extensions of modular Lie algebras*, Proceedings of the International Conference on Algebra, Part 2 (Novosibirsk, 1989) (Providence, RI), Contemp. Math., vol. 131, Amer. Math. Soc., 1992, pp. 31–43. MR 1175819 (93f:17034)

[Gar09] S. Garibaldi, *Vanishing of trace forms in low characteristics*, Algebra Number Theory 3 (2009), no. 5, 543–566, With an appendix by Alexander Premet. MR 2578888 (2011d:20088)

[Gur99] R. M. Guralnick, *Small representations are completely reducible*, J. Algebra 220 (1999), no. 2, 531–541. MR 1717357 (2000m:20018)

[Her13] S. Herpel, *On the smoothness of centralizers in reductive groups*, Trans. Amer. Math. Soc. 365 (2013), no. 7, 3753–3774. MR 3042602

[Hog82] G. M. D. Hogeweij, *Almost-classical Lie algebras. I, II*, Nederl. Akad. Wetensch. Indag. Math. 44 (1982), no. 4, 441–452, 453–460. MR 683531 (84f:17007)

[HS16] Sebastian Herpel and David I. Stewart, *Maximal subalgebras of Cartan type in the exceptional Lie algebras*, Selecta Math. (to appear).

[Hum67] J. E. Humphreys, *Algebraic groups and modular Lie algebras*, Memoirs of the American Mathematical Society, No. 71, American Mathematical Society, Providence, R.I., 1967. MR 0217075 (36 #169)

[Ibr12] S. S. Ibraev, *The second cohomology of simple modules over* $\mathrm{Sp}_4(\mathrm{k})$, Comm. Algebra 40 (2012), no. 3, 1122–1130. MR 2899929

[Jan91] J. C. Jantzen, *First cohomology groups for classical Lie algebras*, Representation theory of finite groups and finite-dimensional algebras (Bielefeld, 1991), Progr. Math., vol. 95, Birkhäuser, Basel, 1991, pp. 289–315. MR 1112165 (92e:17024)

[Jan97] ——, *Low-dimensional representations of reductive groups are semisimple*, Algebraic groups and Lie groups, Austral. Math. Soc. Lect. Ser., vol. 9, Cambridge Univ. Press, Cambridge, 1997, pp. 255–266. MR 1635685 (99g:20079)

[Jan03] ——, *Representations of algebraic groups*, second ed., Mathematical Surveys and Monographs, vol. 107, American Mathematical Society, Providence, RI, 2003. MR MR2015057 (2004h:20061)

[Kan79] W. M. Kantor, *Classical groups from a nonclassical viewpoint*, Oxford University Mathematical Institute, Oxford, 1979. MR 578539 (81h:20056)

[Mar03] Benjamin M. S. Martin, *Reductive subgroups of reductive groups in nonzero characteristic*, J. Algebra 262 (2003), no. 2, 265–286. MR 1971039 (2004g:20066)

[McN02] G. J. McNinch, *The second cohomology of small irreducible modules for simple algebraic groups*, Pacific J. Math. 204 (2002), no. 2, 459–472. MR 1907901 (2003m:20062)

[Mil12] J. S. Milne, *Basic theory of affine group schemes*, available at http://www.jmilne.org/math/CourseNotes/AGS.pdf V.1.00 (2012).

[MT09] George J. McNinch and Donna M. Testerman, *Nilpotent centralizers and Springer isomorphisms*, J. Pure Appl. Algebra 213 (2009), no. 7, 1346–1363. MR 2497582 (2010c:20059)

[Pre95] A. Premet, *Irreducible representations of Lie algebras of reductive groups and the Kac-Weisfeiler conjecture*, Invent. Math. 121 (1995), no. 1, 79–117. MR 1345285 (96g:17007)

[Ric67] R. W. Richardson, Jr., *Conjugacy classes in Lie algebras and algebraic groups*, Ann. of Math. (2) 86 (1967), 1–15. MR 0217079 (36 #173)

[Sei00] Gary M. Seitz, *Unipotent elements, tilting modules, and saturation*, Invent. Math. 141 (2000), no. 3, 467–502. MR 1779618 (2001j:20074)

[Ser94] J-P. Serre, *Sur la semi-simplicité des produits tensoriels de représentations de groupes*, Invent. Math. 116 (1994), no. 1-3, 513–530.

[Ser98] ———, *1998 Moursund lectures at the University of Oregon*, http://arxiv.org/abs/math/0305257v1 (1998).

[SF88] H. Strade and R. Farnsteiner, *Modular Lie algebras and their representations*, Monographs and Textbooks in Pure and Applied Mathematics, vol. 116, Marcel Dekker Inc., New York, 1988. MR 929682 (89h:17021)

[Ste10] D. I. Stewart, *The reductive subgroups of G_2*, J. Group Theory 13 (2010), no. 1, 117–130. MR 2604850 (2011c:20099)

[Ste12] ———, *The second cohomology of simple SL_3-modules*, Comm. Algebra 40 (2012), no. 12, 4702–4716. MR 2989676

[Str73] H. Strade, *Lie algebra representations of dimension $p-1$*, Proc. Amer. Math. Soc. 41 (1973), 419–424. MR 0330247 (48 #8585)

[Ten87] O. K. Ten, *On nonsemisimple maximal subalgebras of Lie p-algebras of classical type*, Vestnik Moskov. Univ. Ser. I Mat. Mekh. (1987), no. 2, 65–67, 103. MR 884515 (88g:17017)

[Tho15] A. R. Thomas, *Simple irreducible subgroups of exceptional algebraic groups*, J. Algebra 423 (2015), 190–238. MR 3283715

[Tho16] ———, *Irreducible A_1 subgroups of exceptional algebraic groups*, J. Algebra 447 (2016), 240–296. MR 3427635

[Vas05] A. Vasiu, *Normal, unipotent subgroup schemes of reductive groups*, C. R. Math. Acad. Sci. Paris 341 (2005), no. 2, 79–84. MR 2153960 (2006g:14076)

[Wri11] C. B. Wright, *Second cohomology groups for algebraic groups and their Frobenius kernels*, Journal of Algebra 330 (2011), no. 1, 60 – 75.

Sebastian Herpel
Ruhr-Universität Bochum
Bochum, Germany

David Stewart
University of Cambridge
Cambridge, UK
dis20@cam.ac.uk

Milne's Correcting Factor and Derived De Rham Cohomology

Baptiste Morin

Received: June 15, 2015
Revised: November 13, 2015

Communicated by Stephen Lichtenbaum

Abstract. Milne's correcting factor is a numerical invariant playing an important role in formulas for special values of zeta functions of varieties over finite fields. We show that Milne's factor is simply the Euler characteristic of the derived de Rham complex (relative to $\mathbb{Z}$) modulo the Hodge filtration.

2010 Mathematics Subject Classification: 14G10, 14F40, 11S40, 11G25
Keywords and Phrases: Zeta functions, Special values, Derived de Rham cohomology

A result of Milne ([9] Theorem 0.1) describes the special values of the zeta function of a smooth projective variety X over a finite field satisfying the Tate conjecture. A very natural reformulation of this result was given by Lichtenbaum and Geisser (see [2], [7], [8] and [10]) using Weil-étale cohomology of motivic complexes. They conjecture that

$$\lim_{t\to q^{-n}} Z(X,t)\cdot(1-q^n t)^{\rho_n} = \pm\chi(H^*_W(X,\mathbb{Z}(n)),\cup e)\cdot q^{\chi(X/\mathbb{F}_q,\mathcal{O}_X,n)} \tag{1}$$

and show that (1) holds whenever the groups $H^i_W(X,\mathbb{Z}(n))$ are finitely generated. Here $H^*_W(X,\mathbb{Z}(n))$ denotes Weil-étale motivic cohomology, $e\in H^1(W_{\mathbb{F}_q},\mathbb{Z})$ is a fundamental class and $\chi(H^*_W(X,\mathbb{Z}(n)),e)$ is the Euler characteristic of the complex

$$\cdots\xrightarrow{\cup e} H^i_W(X,\mathbb{Z}(n))\xrightarrow{\cup e} H^{i+1}_W(X,\mathbb{Z}(n))\xrightarrow{\cup e}\cdots \tag{2}$$

More precisely, the cohomology groups of the complex (2) are finite and $\chi(H^*_W(X,\mathbb{Z}(n)),\cup e)$ is the alternating product of their orders. Finally, Milne's correcting factor $q^{\chi(X/\mathbb{F}_q,\mathcal{O},n)}$ was defined in [9] by the formula

$$\chi(X/\mathbb{F}_q,\mathcal{O}_X,n)=\sum_{i\le n,j}(-1)^{i+j}\cdot(n-i)\cdot\dim_{\mathbb{F}_q}H^j(X,\Omega^i_{X/\mathbb{F}_q}).$$

The author was supported by ANR-12-BS01-0002 and ANR-12-JS01-0007.

It is possible to generalize (1) in order to give a conjectural description of special values of zeta functions of all separated schemes of finite type over $\mathbb{F}_q$ (see [3] Conjecture 1.4), and even of all motivic complexes over $\mathbb{F}_q$ (see [11] Conjecture 1.2). The statement of those more general conjectures is in any case very similar to formula (1). The present note is motivated by the hope for a further generalization, which would apply to zeta functions of all algebraic schemes over Spec($\mathbb{Z}$). As briefly explained below, the special-value conjecture for (flat) schemes over Spec($\mathbb{Z}$) must take a rather different form than formula (1). Going back to the special case of smooth projective varieties over finite fields, this leads to a slightly different restatement of formula (1).
Let $\mathcal{X}$ be a regular scheme proper over Spec($\mathbb{Z}$). The "fundamental line"

$$\Delta(\mathcal{X}/\mathbb{Z},n) := \det_{\mathbb{Z}} R\Gamma_{W,c}(\mathcal{X},\mathbb{Z}(n)) \otimes_{\mathbb{Z}} \det_{\mathbb{Z}} R\Gamma_{dR}(\mathcal{X}/\mathbb{Z})/F^n$$

should be a well defined invertible $\mathbb{Z}$-module endowed with a canonical trivialization

$$\mathbb{R} \xrightarrow{\sim} \Delta(\mathcal{X}/\mathbb{Z},n) \otimes_{\mathbb{Z}} \mathbb{R}.$$

involving a fundamental class $\theta \in H^1(\mathbb{R},\mathbb{R}) =$ "$H^1(W_{\mathbb{F}_1},\mathbb{R})$" analogous to $e \in H^1(W_{\mathbb{F}_q},\mathbb{Z})$. Here $R\Gamma_{W,c}(\mathcal{X},\mathbb{Z}(n))$ denotes Weil-étale cohomology with compact support. However, there is no natural trivialization $\mathbb{R} \xrightarrow{\sim} \det_{\mathbb{Z}} R\Gamma_{W,c}(\mathcal{X},\mathbb{Z}(n)) \otimes_{\mathbb{Z}} \mathbb{R}$. Consequently, it is not possible to define an Euler characteristic generalizing $\chi(H^*_W(X,\mathbb{Z}(n)),\cup e)$, neither to define a correcting factor generalizing Milne's correcting factor: one is forced to consider the fundamental line as a whole. Let us go back to the case of smooth projective varieties $X/\mathbb{F}_q$, which we now see as schemes over $\mathbb{Z}$. Accordingly, we replace $Z(X,t)$ with $\zeta(X,s) = Z(X,q^{-s})$, the fundamental class e with θ and the cotangent sheaf $\Omega^1_{X/\mathbb{F}_q} \simeq L_{X/\mathbb{F}_q}$ with the cotangent complex $L_{X/\mathbb{Z}}$. Assuming that $H^i_W(X,\mathbb{Z}(n))$ is finitely generated for all i, the fundamental line

$$\Delta(X/\mathbb{Z},n) := \det_{\mathbb{Z}} R\Gamma_W(X,\mathbb{Z}(n)) \otimes_{\mathbb{Z}} \det_{\mathbb{Z}} R\Gamma(X, L\Omega^*_{X/\mathbb{Z}}/F^n) \tag{3}$$

is well defined and cup-product with θ gives a trivialization

$$\lambda : \mathbb{R} \xrightarrow{\sim} \Delta(X/\mathbb{Z},n) \otimes_{\mathbb{Z}} \mathbb{R}.$$

Here $L\Omega^*_{X/\mathbb{Z}}/F^n$ is Illusie's derived de Rham complex modulo the Hodge filtration (see [6] VIII.2.1). The aim of this note is to show that the Euler characteristic of $R\Gamma(X, L\Omega^*_{X/\mathbb{Z}}/F^n)$ equals $q^{\chi(X/\mathbb{F}_q,\mathcal{O}_X,n)}$, hence that Milne's correcting factor is naturally part of the fundamental line. We denote by $\zeta^*(X,n)$ the leading coefficient in the Taylor development of $\zeta(X,s)$ near $s=n$.

THEOREM. *Let X be a smooth proper scheme over $\mathbb{F}_q$ and let $n \in \mathbb{Z}$ be an integer. Then we have*

$$\prod_{i\in\mathbb{Z}} \mid H^i(X, L\Omega^*_{X/\mathbb{Z}}/F^n) \mid^{(-1)^i} \;=\; q^{\chi(X/\mathbb{F}_q,\mathcal{O}_X,n)}.$$

Assume moreover that X is projective and that the groups $H^i_W(X,\mathbb{Z}(n))$ are finitely generated for all i. Then one has

$$\begin{aligned}\Delta(X/\mathbb{Z},n) &= \mathbb{Z}\cdot\lambda\left(\log(q)^{\rho_n}\cdot\chi(H^*_W(X,\mathbb{Z}(n)),\cup e)^{-1}\cdot q^{-\chi(X/\mathbb{F}_q,\mathcal{O}_X,n)}\right)\\ &= \mathbb{Z}\cdot\lambda\left(\zeta^*(X,n)^{-1}\right)\end{aligned}$$

where $\rho_n := -\mathrm{ord}_{s=n}\zeta(X,s)$ is the order of the pole of $\zeta(X,s)$ at $s=n$.

Before giving the proof, we need to fix some notations. For an object C in the derived category of abelian groups such that $H^i(C)$ is finitely generated for all i and $H^i(C)=0$ for almost all i, we set

$$\det_{\mathbb{Z}}(C) := \bigotimes_{i\in\mathbb{Z}} \det_{\mathbb{Z}}^{(-1)^i} H^i(C).$$

If $H^i(C)$ is moreover finite for all i, then we call the following isomorphism

$$\det_{\mathbb{Z}}(C)\otimes_{\mathbb{Z}}\mathbb{Q} \xrightarrow{\sim} \bigotimes_{i\in\mathbb{Z}} \det_{\mathbb{Q}}^{(-1)^i}\left(H^i(C)\otimes_{\mathbb{Z}}\mathbb{Q}\right) \xrightarrow{\sim} \bigotimes_{i\in\mathbb{Z}} \det_{\mathbb{Q}}^{(-1)^i}(0) \xrightarrow{\sim} \mathbb{Q}$$

the *canonical* $\mathbb{Q}$*-trivialization* of $\det_{\mathbb{Z}}(C)$. Let A be a finite abelian group, which we see as a complex concentrated in degree 0. Then the canonical $\mathbb{Q}$-trivialization $\det_{\mathbb{Z}}(A)\otimes_{\mathbb{Z}}\mathbb{Q}\simeq\mathbb{Q}$ identifies $\det_{\mathbb{Z}}(A)$ with $|A|^{-1}\cdot\mathbb{Z}\subset\mathbb{Q}$, where $|A|$ denotes the order of A.
Given a ring R and an R-module M, we denote by $\Gamma_R(M)$ the universal divided power R-algebra of M, and by $\Gamma^i_R(M)$ its submodule of homogeneous elements of degree i. We refer to ([1] Appendix A) for the definition of $\Gamma_R(M)$ and its main properties. There is a canonical map $\gamma^i: M\to\Gamma^i_R(M)$, such that composition with γ^i induces a bijection $\mathrm{Hom}_R(\Gamma^i_R(M),N)\xrightarrow{\sim}P^i(M,N)$, where $P^i(M,N)$ is the set of "homogeneous polynomial functions of degree i". The functor Γ^i_R sends free modules of finite type to free modules of finite type. Moreover Γ^i_R commutes with filtered colimits, hence sends flat modules to flat modules. If M is free of rank one, then so is $\Gamma^i_R(M)$. If (T,R) is a ringed topos and M an R-module, then $\Gamma_R(M)$ is the sheafification of $U\mapsto\Gamma_{R(U)}(M(U))$. We also denote by Λ^i_R the (non-additive) exterior power functor and by $L\Lambda^i_R$ its left derived functor (see [5] I.4.2). We often omit the subscript R and simply write $\Gamma^i M$, $\Lambda^i M$ and $L\Lambda^i M$.
Let X be a scheme. The notation $R\Gamma(X,-)$ refers to hypercohomology with respect to the Zariski topology.

Proof. Since Milne's correcting factor is insensitive to restriction of scalars (i.e. $q^{\chi(X/\mathbb{F}_q,\mathcal{O}_X,n)} = p^{\chi(X/\mathbb{F}_p,\mathcal{O}_X,n)}$), we may consider X over $\mathbb{F}_p$. We need the following

Lemma 1. *Let $E^{*,*}_* = (E^{p,q}_r, d^{p,q}_r)^{p,q}_r$ be a cohomological spectral sequence of abelian groups with abutment H^*. Assume that there exists an index r_0 such that $E^{p,q}_{r_0}$ is finite for all $(p,q)\in\mathbb{Z}^2$ and $E^{p,q}_{r_0}=0$ for all but finitely many*

(p,q). Then we have a canonical isomorphism

$$\iota : \bigotimes_{p,q} \det_{\mathbb{Z}}^{(-1)^{p+q}} E_{r_0}^{p,q} \xrightarrow{\sim} \bigotimes_{n} \det_{\mathbb{Z}}^{(-1)^n} H^n$$

such that the square of isomorphisms

$$\begin{tikzcd} \left(\bigotimes_{p,q} \det_{\mathbb{Z}}^{(-1)^{p+q}} E_{r_0}^{p,q}\right) \otimes \mathbb{Q} \arrow[r, "\iota\otimes\mathbb{Q}"] \arrow[d] & \left(\bigotimes_{n} \det_{\mathbb{Z}}^{(-1)^n} H^n\right) \otimes \mathbb{Q} \arrow[d] \\ \mathbb{Q} \arrow[r, "\mathrm{Id}"] & \mathbb{Q} \end{tikzcd}$$

commutes, where the vertical maps are the canonical $\mathbb{Q}$-trivializations.

Proof. For any $t \geq r_0$, consider the bounded cochain complex C_t^* of finite abelian groups:

$$\cdots \longrightarrow \bigoplus_{p+q=n-1} E_t^{p,q} \longrightarrow \bigoplus_{p+q=n} E_t^{p,q} \xrightarrow{\oplus d_t^{p,q}} \bigoplus_{p+q=n+1} E_t^{p,q} \longrightarrow \cdots$$

The fact that the cohomology of C_t^* is given by $H^n(C_t^*) = \bigoplus_{p+q=n} E_{t+1}^{p,q}$ gives an isomorphism

$$\bigotimes_{p,q} \det_{\mathbb{Z}}^{(-1)^{p+q}} E_t^{p,q} \xrightarrow{\sim} \bigotimes_{p,q} \det_{\mathbb{Z}}^{(-1)^{p+q}} E_{t+1}^{p,q}$$

compatible with the canonical $\mathbb{Q}$-trivializations. By assumption, there exists an index $r_1 \geq r_0$ such that the spectral sequence degenerates at the r_1-page, i.e. $E_{r_1}^{*,*} = E_{\infty}^{*,*}$. The induced filtration on each H^n is such that $\mathrm{gr}^p H^n = E_{\infty}^{p,n-p}$. We obtain isomorphisms

$$\bigotimes_{p,q} \det_{\mathbb{Z}}^{(-1)^{p+q}} E_{r_0}^{p,q} \xrightarrow{\sim} \bigotimes_{p,q} \det_{\mathbb{Z}}^{(-1)^{p+q}} E_{\infty}^{p,q} \xrightarrow{\sim}$$
$$\xrightarrow{\sim} \bigotimes_{n} \bigotimes_{p} \det_{\mathbb{Z}}^{(-1)^n} E_{\infty}^{p,n-p} \xrightarrow{\sim} \bigotimes_{n} \det_{\mathbb{Z}}^{(-1)^n} H^n$$

compatible with the canonical $\mathbb{Q}$-trivializations. □

Consider the Hodge filtration F^* on the derived de Rham complex $L\Omega_{X/\mathbb{Z}}^*$. By ([6] VIII.2.1.1.5) we have

$$\mathrm{gr}(L\Omega_{X/\mathbb{Z}}^*) \simeq \bigoplus_{p\geq 0} L\Lambda^p L_{X/\mathbb{Z}}[-p].$$

This gives a (convergent) spectral sequence

$$E_1^{p,q} = H^q(X, L\Lambda^{p<n} L_{X/\mathbb{Z}}) \Longrightarrow H^{p+q}(X, L\Omega_{X/\mathbb{Z}}^*/F^n)$$

where $L\Lambda^{p<n} L_{X/\mathbb{Z}} := L\Lambda^p L_{X/\mathbb{Z}}$ for $p < n$ and $L\Lambda^{p<n} L_{X/\mathbb{Z}} := 0$ otherwise. The scheme X is proper and $L\Lambda^p L_{X/\mathbb{Z}}$ is isomorphic, in the derived category $\mathcal{D}(\mathcal{O}_X)$ of $\mathcal{O}_X$-modules, to a bounded complex of coherent sheaves (see (6) below). It

follows that $E_1^{p,q}$ is a finite dimensional $\mathbb{F}_p$-vector space for all (p,q) vanishing for almost all (p,q). By Lemma 1, this yields isomorphisms

$$\begin{aligned}\det_{\mathbb{Z}} R\Gamma(X, L\Omega^*_{X/\mathbb{Z}}/F^n) &\xrightarrow{\sim} \bigotimes_i \det_{\mathbb{Z}}^{(-1)^i} H^i(X, L\Omega^*_{X/\mathbb{Z}}/F^n)\\ &\xrightarrow{\sim} \bigotimes_{p<n,q} \det_{\mathbb{Z}}^{(-1)^{p+q}} H^q(X, L\Lambda^p L_{X/\mathbb{Z}})\\ &\xrightarrow{\sim} \bigotimes_{p<n} \det_{\mathbb{Z}}^{(-1)^p} R\Gamma(X, L\Lambda^p L_{X/\mathbb{Z}})\end{aligned}$$

which are compatible with the canonical $\mathbb{Q}$-trivializations. The transitivity triangle (see [5] II.2.1) for the composite map $X \xrightarrow{f} \mathrm{Spec}(\mathbb{F}_p) \to \mathrm{Spec}(\mathbb{Z})$ reads as follows (using [5] III.3.1.2 and [5] III.3.2.4(iii)):

$$Lf^*(p\mathbb{Z}/p^2\mathbb{Z})[1] \to L_{X/\mathbb{Z}} \to \Omega^1_{X/\mathbb{F}_p}[0] \xrightarrow{\omega} Lf^*(p\mathbb{Z}/p^2\mathbb{Z})[2]. \tag{4}$$

We set $\mathcal{L} := Lf^*(p\mathbb{Z}/p^2\mathbb{Z})$, a trivial invertible $\mathcal{O}_X$-module. By ([5] Théorème III.2.1.7), the class

$$\omega \in \mathrm{Ext}^2_{\mathcal{O}_X}(\Omega^1_{X/\mathbb{F}_p}, \mathcal{L}) \simeq H^2(X, T_{X/\mathbb{F}_p})$$

is the obstruction to the existence of a lifting of X over $\mathbb{Z}/p^2\mathbb{Z}$. If such a lifting does exist then we have $\omega = 0$, in which case the following lemma is superfluous. For an object C of $\mathcal{D}(\mathcal{O}_X)$ with bounded cohomology, we set

$$\mathrm{gr}_\tau C := \bigoplus_{i\in\mathbb{Z}} H^i(C)[-i].$$

Lemma 2. *We have an isomorphism*

$$\det_{\mathbb{Z}} R\Gamma(X, L\Lambda^p L_{X/\mathbb{Z}}) \simeq \det_{\mathbb{Z}} R\Gamma(X, L\Lambda^p(\mathrm{gr}_\tau L_{X/\mathbb{Z}}))$$

compatible with the canonical $\mathbb{Q}$-trivializations.

Proof. The map $X \to \mathrm{Spec}(\mathbb{Z})$ is a local complete intersection, hence the complex $L_{X/\mathbb{Z}}$ has perfect amplitude $\subset [-1,0]$ (see [5] III.3.2.6). In other words, $L_{X/\mathbb{Z}}$ is locally isomorphic in $\mathcal{D}(\mathcal{O}_X)$ to a complex of free modules of finite type concentrated in degrees -1 and 0. By ([4] 2.2.7.1) and ([4] 2.2.8), $L_{X/\mathbb{Z}}$ is globally isomorphic to such a complex, i.e. there exists an isomorphism $L_{X/\mathbb{Z}} \simeq [M \to N]$ in $\mathcal{D}(\mathcal{O}_X)$, where M and N are finitely generated locally free $\mathcal{O}_X$-modules put in degrees -1 and 0 respectively. Consider the exact sequences

$$0 \to \mathcal{L} \to M \to F \to 0 \text{ and } 0 \to F \to N \to \Omega \to 0 \tag{5}$$

where $\mathcal{L} := Lf^*(p\mathbb{Z}/p^2\mathbb{Z})$ and $\Omega := \Omega^1_{X/\mathbb{F}_p}$ are finitely generated and locally free. It follows that F is also finitely generated and locally free. One has an isomorphism in $\mathcal{D}(\mathcal{O}_X)$

$$L\Lambda^p L_{X/\mathbb{Z}} \simeq [\Gamma^p M \to \Gamma^{p-1} M \otimes N \to \cdots \to M \otimes \Lambda^{p-1} N \to \Lambda^p N] \tag{6}$$

where the right hand side sits in degrees $[-p,0]$ (see [6] VIII.2.1.2 and [5] I.4.3.2.1). Moreover, in view of (4) we may choose an isomorphism

$$\mathrm{gr}_{\tau} L_{X/\mathbb{Z}} \simeq [\mathcal{L} \xrightarrow{0} \Omega]$$

in $\mathcal{D}(\mathcal{O}_X)$, the right hand side being concentrated in degrees $[-1,0]$. Hence the complex $L\Lambda^p(\mathrm{gr}_{\tau} L_{X/\mathbb{Z}}) \in \mathcal{D}(\mathcal{O}_X)$ is represented by a complex of the form

$$\begin{aligned} (7)\quad L\Lambda^p(\mathrm{gr}_{\tau} L_{X/\mathbb{Z}}) &\simeq L\Lambda^p([\mathcal{L} \to \Omega]) \simeq \\ &\simeq [\Gamma^p \mathcal{L} \to \Gamma^{p-1}\mathcal{L} \otimes \Omega \to \cdots \to \mathcal{L} \otimes \Lambda^{p-1}\Omega \to \Lambda^p \Omega] \end{aligned}$$

where the right hand side sits in degrees $[-p,0]$. Lemma 1 and (6) give an isomorphism

$$(8)\qquad \det_{\mathbb{Z}} R\Gamma(X, L\Lambda^p L_{X/\mathbb{Z}}) \simeq \bigotimes_{0\leq q\leq p} \det_{\mathbb{Z}}^{(-1)^{p-q}} R\Gamma(X, \Gamma^{p-q}M \otimes \Lambda^q N)$$

compatible with the $\mathbb{Q}$-trivializations. The second exact sequence in (5) endows $\Lambda^q N$ with a finite decreasing filtration Fil^* such that $\mathrm{gr}^i_{\mathrm{Fil}}(\Lambda^q N) = \Lambda^i F \otimes \Lambda^{q-i}\Omega$. Since $\Gamma^{p-q}M$ is flat, Fil^* induces a similar filtration on $\Gamma^{p-q}M \otimes \Lambda^q N$ such that

$$\mathrm{gr}^i_{\mathrm{Fil}}(\Gamma^{p-q}M \otimes \Lambda^q N) = \Gamma^{p-q}M \otimes \Lambda^i F \otimes \Lambda^{q-i}\Omega.$$

This filtration induces an isomorphism

$$(9)\quad \det_{\mathbb{Z}} R\Gamma(X, \Gamma^{p-q}M \otimes \Lambda^q N) \simeq \bigotimes_{0\leq i\leq q} \det_{\mathbb{Z}} R\Gamma(X, \Gamma^{p-q}M \otimes \Lambda^i F \otimes \Lambda^{q-i}\Omega)$$

compatible with the $\mathbb{Q}$-trivializations. Lemma 1 and (7) give an isomorphism

$$(10)\qquad \det_{\mathbb{Z}} R\Gamma(X, L\Lambda^p(\mathrm{gr}_{\tau} L_{X/\mathbb{Z}})) \simeq \bigotimes_{0\leq i\leq p} \det_{\mathbb{Z}}^{(-1)^{p-i}} R\Gamma(X, \Gamma^{p-i}\mathcal{L} \otimes \Lambda^i \Omega)$$

compatible with the $\mathbb{Q}$-trivializations. Moreover, we have an isomorphism (see [5] I.4.3.1.7)

$$\Gamma^{p-i}\mathcal{L} \simeq [\Gamma^{p-i}M \to \Gamma^{p-i-1}M \otimes F \to \cdots \to M \otimes \Lambda^{p-i-1}F \to \Lambda^{p-i}F]$$

where the right hand side sits in degrees $[0, p-i]$. Since $\Lambda^i\Omega$ is flat, we have an isomorphism between $\Gamma^{p-i}\mathcal{L} \otimes \Lambda^i\Omega$ and

$$\begin{aligned} &[\Gamma^{p-i}M \otimes \Lambda^i\Omega \to \Gamma^{p-i-1}M \otimes F \otimes \Lambda^i\Omega \to \cdots \\ &\qquad \cdots \to M \otimes \Lambda^{p-i-1}F \otimes \Lambda^i\Omega \to \Lambda^{p-i}F \otimes \Lambda^i\Omega]. \end{aligned}$$

By Lemma 1, we have
(11)

$$\det_{\mathbb{Z}} R\Gamma(X, \Gamma^{p-i}\mathcal{L} \otimes \Lambda^i\Omega) \simeq \bigotimes_{0\leq j\leq p-i} \det_{\mathbb{Z}}^{(-1)^j} R\Gamma(X, \Gamma^{p-i-j}M \otimes \Lambda^j F \otimes \Lambda^i\Omega).$$

Putting (10), (11), (9) and (8) together, we obtain isomorphisms

$$\begin{aligned}
&\det_{\mathbb{Z}} R\Gamma(X, L\Lambda^p(\mathrm{gr}_\tau L_{X/\mathbb{Z}})) \simeq \\
&\simeq \bigotimes_{0\leq i\leq p} \det_{\mathbb{Z}}^{(-1)^{p-i}} R\Gamma(X, \Gamma^{p-i}\mathcal{L}\otimes\Lambda^i\Omega) \\
&\simeq \bigotimes_{0\leq i\leq p}\left(\bigotimes_{0\leq j\leq p-i} \det_{\mathbb{Z}}^{(-1)^{p-i-j}} R\Gamma(X, \Gamma^{p-i-j}M\otimes\Lambda^j F\otimes\Lambda^i\Omega)\right) \\
&= \bigotimes_{0\leq q\leq p}\left(\bigotimes_{0\leq i,j\ ;\, i+j=q} \det_{\mathbb{Z}}^{(-1)^{p-q}} R\Gamma(X, \Gamma^{p-q}M\otimes\Lambda^j F\otimes\Lambda^i\Omega)\right) \\
&\simeq \bigotimes_{0\leq q\leq p} \det_{\mathbb{Z}}^{(-1)^{p-q}} R\Gamma(X, \Gamma^{p-q}M\otimes\Lambda^q N) \\
&\simeq \det_{\mathbb{Z}} R\Gamma(X, L\Lambda^p L_{X/\mathbb{Z}})
\end{aligned}$$

compatible with the canonical $\mathbb{Q}$-trivializations. $\square$

Recall from (7) that the complex $L\Lambda^p(\mathrm{gr}_\tau L_{X/\mathbb{Z}})$ is isomorphic in $\mathcal{D}(\mathcal{O}_X)$ to a complex of the form

$$0\to\Gamma^p\mathcal{L}\to\Gamma^{p-1}\mathcal{L}\otimes\Omega^1_{X/\mathbb{F}_p}\to\cdots\to\Gamma^1\mathcal{L}\otimes\Omega^{p-1}_{X/\mathbb{F}_p}\to\Omega^p_{X/\mathbb{F}_p}\to 0$$

put in degrees $[-p,0]$. An isomorphism of $\mathbb{F}_p$-vector spaces $\mathbb{F}_p\simeq p\mathbb{Z}/p^2\mathbb{Z}$ induces an identification $\mathcal{O}_X\simeq\mathcal{L}$, and more generally $\mathcal{O}_X\simeq\Gamma^i\mathcal{L}$ for any $i\geq 0$. Hence $(L\Lambda^p(\mathrm{gr}_\tau L_{X/\mathbb{Z}}))[-p]\in\mathcal{D}(\mathcal{O}_X)$ is represented by a complex of the form

$$0\to\mathcal{O}_X\to\Omega^1_{X/\mathbb{F}_p}\to\cdots\to\Omega^p_{X/\mathbb{F}_p}\to 0 \tag{12}$$

sitting in degrees $[0,p]$. We obtain a spectral sequence

$$E_1^{i,j}=H^j(X,\Omega^{i\leq p}_{X/\mathbb{F}_p})\Longrightarrow H^{i+j}(X,(L\Lambda^p(\mathrm{gr}_\tau L_{X/\mathbb{Z}}))[-p])$$

where $\Omega^{i\leq p}:=\Omega^i$ for $i\leq p$ and $\Omega^{i\leq p}:=0$ for $i>p$. By Lemma 1 again, we get an identification

$$\begin{aligned}
\bigotimes_{i\leq p,j}\det_{\mathbb{Z}}^{(-1)^{i+j}} H^j(X,\Omega^i_{X/\mathbb{F}_p}) &\xrightarrow{\sim} \det_{\mathbb{Z}} R\Gamma(X,(L\Lambda^p(\mathrm{gr}_\tau L_{X/\mathbb{Z}}))[-p]) \\
&\xrightarrow{\sim} \det_{\mathbb{Z}}^{(-1)^p} R\Gamma(X, L\Lambda^p(\mathrm{gr}_\tau L_{X/\mathbb{Z}})).
\end{aligned}$$

In summary, we have the following isomorphisms

$$\det_{\mathbb{Z}} R\Gamma(X, L\Omega^*_{X/\mathbb{Z}}/F^n) \xrightarrow{\sim} \bigotimes_{p<n}\det_{\mathbb{Z}}^{(-1)^p} R\Gamma(X, L\Lambda^p L_{X/\mathbb{Z}}) \tag{13}$$

$$\xrightarrow{\sim} \bigotimes_{p<n}\det_{\mathbb{Z}}^{(-1)^p} R\Gamma(X, L\Lambda^p(\mathrm{gr}_\tau L_{X/\mathbb{Z}})) \tag{14}$$

$$\xrightarrow{\sim} \bigotimes_{p<n}\left(\bigotimes_{i\leq p,j}\det_{\mathbb{Z}}^{(-1)^{i+j}} H^j(X,\Omega^i_{X/\mathbb{F}_p})\right) \tag{15}$$

such that the square

$$\begin{tikzcd}
\left(\det_{\mathbb{Z}} R\Gamma(X, L\Omega^*_{X/\mathbb{Z}}/F^n)\right)\otimes\mathbb{Q} \arrow[r] \arrow[d, "\gamma"] & \left(\bigotimes_{p<n}\bigotimes_{i\leq p,j}\det_{\mathbb{Z}}^{(-1)^{i+j}} H^j(X,\Omega^i_{X/\mathbb{F}_p})\right)\otimes\mathbb{Q} \arrow[d, "\gamma'"] \\
\mathbb{Q} \arrow[r, "\mathrm{Id}"] & \mathbb{Q}
\end{tikzcd}$$

commutes, where the top horizontal map is induced by (15), and the vertical isomorphisms are the canonical trivializations. The first assertion of the theorem follows:

$$\begin{aligned}
\mathbb{Z}\cdot\left(\prod_{i\in\mathbb{Z}} \mid H^i(X, L\Omega^*_{X/\mathbb{Z}}/F^n)\mid^{(-1)^i}\right)^{-1} &= \\
&= \gamma\left(\det_{\mathbb{Z}} R\Gamma(X, L\Omega^*_{X/\mathbb{Z}}/F^n)\right)\\
&= \gamma'\left(\bigotimes_{p<n}\bigotimes_{i\leq p,j}\det_{\mathbb{Z}}^{(-1)^{i+j}} H^j(X,\Omega^i_{X/\mathbb{F}_p})\right)\\
&= \mathbb{Z}\cdot p^{-\chi(X/\mathbb{F}_p,\mathcal{O}_X,n)}.
\end{aligned}$$

We now explain why the second assertion of the theorem is a restatement of ([2] Theorem 1.3). We assume that $H^i_W(X,\mathbb{Z}(n))$ is finitely generated for all $i\in\mathbb{Z}$ (X and n being fixed). Recall from [2] that this assumption implies the following: $H^i_W(X,\mathbb{Z}(n))$ is in fact finite for $i\neq 2n, 2n+1$, the complex (2) has finite cohomology groups and one has

$$\rho_n := -\mathrm{ord}_{s=n}\zeta(X,s) = \mathrm{rank}_{\mathbb{Z}} H^{2n}_W(X,\mathbb{Z}(n)).$$

In particular the complex

$$\cdots \xrightarrow{\cup e} H^i_W(X,\mathbb{Z}(n))\otimes\mathbb{Q} \xrightarrow{\cup e} H^{i+1}_W(X,\mathbb{Z}(n))\otimes\mathbb{Q} \xrightarrow{\cup e} \cdots \tag{16}$$

is acyclic. This gives a trivialization

$$\begin{aligned}
\beta:\mathbb{Q} &\xrightarrow{\sim} \bigotimes_i \det_{\mathbb{Q}}^{(-1)^i}\left(H^i_W(X,\mathbb{Z}(n))\otimes\mathbb{Q}\right) \xrightarrow{\sim}\\
&\xrightarrow{\sim} \left(\bigotimes_i \det_{\mathbb{Z}}^{(-1)^i} H^i_W(X,\mathbb{Z}(n))\right)\otimes\mathbb{Q}
\end{aligned}$$

such that

$$\mathbb{Z}\cdot\beta\left(\chi(H^*_W(X,\mathbb{Z}(n)),\cup e)^{-1}\right) = \bigotimes_i \det_{\mathbb{Z}}^{(-1)^i} H^i_W(X,\mathbb{Z}(n)).$$

The class $e \in H^1(W_{\mathbb{F}_q}, \mathbb{Z}) = \mathrm{Hom}(W_{\mathbb{F}_q}, \mathbb{Z})$ maps the Frobenius Frob $\in W_{\mathbb{F}_q}$ to $1 \in \mathbb{Z}$. We define the map

$$W_{\mathbb{F}_q} = \mathbb{Z} \cdot \mathrm{Frob} \longrightarrow \mathbb{R} =: W_{\mathbb{F}_1}$$

as the map sending Frob to $\log(q)$, while $\theta \in H^1(W_{\mathbb{F}_1}, \mathbb{R}) = \mathrm{Hom}(\mathbb{R}, \mathbb{R})$ is the identity map. It follows that the acyclic complex

$$\cdots \xrightarrow{\cup\theta} H^i_W(X, \mathbb{Z}(n)) \otimes \mathbb{R} \xrightarrow{\cup\theta} H^{i+1}_W(X, \mathbb{Z}(n)) \otimes \mathbb{R} \xrightarrow{\cup\theta} \cdots$$

induces a trivialization

$$\alpha : \mathbb{R} \xrightarrow{\sim} \bigotimes_i \det\nolimits_{\mathbb{R}}^{(-1)^i} \left(H^i_W(X, \mathbb{Z}(n)) \otimes \mathbb{R}\right) \xrightarrow{\sim} \left(\bigotimes_i \det\nolimits_{\mathbb{Z}}^{(-1)^i} H^i_W(X, \mathbb{Z}(n))\right) \otimes \mathbb{R}$$

such that

$$\mathbb{Z} \cdot \alpha \left(\chi(H^*_W(X, \mathbb{Z}(n)), \cup e)^{-1} \cdot \log(q)^{\rho_n}\right) = \bigotimes_i \det\nolimits_{\mathbb{Z}}^{(-1)^i} H^i_W(X, \mathbb{Z}(n)).$$

The trivialization λ is the product of α with the canonical trivialization

$$\mathbb{R} \xrightarrow{\sim} \det\nolimits_{\mathbb{Z}} R\Gamma(X, L\Omega^*_{X/\mathbb{Z}}/F^n) \otimes_{\mathbb{Z}} \mathbb{R}.$$

Hence we have

$$\mathbb{Z} \cdot \lambda \left(\log(q)^{\rho_n} \cdot \chi(H^*_W(X, \mathbb{Z}(n)), \cup e)^{-1} \cdot q^{-\chi(X, \mathcal{O}_X, n)}\right) = \Delta(X/\mathbb{Z}, n).$$

Moreover, formula (1) gives

$$\zeta^*(X, s) = \pm \log(q)^{-\rho_n} \cdot \chi(H^*_W(X, \mathbb{Z}(n)), \cup e) \cdot q^{\chi(X, \mathcal{O}_X, n)}$$

hence the result follows from ([2] Theorem 1.3). □

Acknowledgments. I would like to thank Matthias Flach, Stephen Lichtenbaum and Niranjan Ramachandran for their interest and comments.

References

[1] Berthelot, P.; Ogus, A.: *Notes on crystalline cohomology.* Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1978.

[2] Geisser, T.: *Weil-étale cohomology over finite fields.* Math. Ann. 330 (4) (2004), 665–692.

[3] Geisser, T.: *Arithmetic cohomology over finite fields and special valyues of ζ-functions.* Duke Math. J. 133 (1) (2006), 27–57.

[4] Illusie, L.: *Existence de Résolutions Globales.* Théorie des Intersections et Théorème de Riemann-Roch (SGA6), Lecture Notes in Mathematics 225, (1971), 160–221.

[5] Illusie, L.: *Complexe cotangent et déformations. I.* Lecture Notes in Mathematics, Vol. 239. Springer-Verlag, Berlin-New York, 1971.

[6] Illusie, L.: *Complexe cotangent et déformations. II.* Lecture Notes in Mathematics, Vol. 283. Springer-Verlag, Berlin-New York, 1972.

[7] Lichtenbaum, S.: *Values of zeta-functions at nonnegative integers.* In: Number theory, Noordwijkerhout 1983, Lecture Notes in Mathematics, Vol. 1068, 127–138, Springer, Berlin, 1984.

[8] Lichtenbaum, S.: *The Weil-étale topology on schemes over finite fields.* Compositio Math. 141 (3) (2005), 689–702.

[9] Milne, J.S.: *Values of zeta functions of varieties over finite fields.* Amer. J. Math. 108 (1986), no. 2, 297–360.

[10] Milne, J.S.: *Addendum to: Values of zeta functions of varieties over finite fields, Amer. J. Math. 108, (1986), 297-360.* (2013), available at http://www.jmilne.org/.

[11] Milne, J.S., Ramachandran, N.: *Motivic complexes and special values of zeta functions.* Preprint 2013, arXiv:1311.3166.

Baptiste Morin
CNRS, IMB
Université de Bordeaux
351, cours de la Libération
F 33405 Talence cedex,
France
Baptiste.Morin@math.u-bordeaux1.fr

On Additive Higher Chow Groups of Affine Schemes

Amalendu Krishna and Jinhyun Park

Received: July 5, 2015
Revised: December 18, 2015

Communicated by Takeshi Saito

Abstract. We show that the multivariate additive higher Chow groups of a smooth affine k-scheme $\mathrm{Spec}\,(R)$ essentially of finite type over a perfect field k of characteristic $\neq 2$ form a differential graded module over the big de Rham-Witt complex $\mathbb{W}_m\Omega^\bullet_R$. In the univariate case, we show that additive higher Chow groups of $\mathrm{Spec}\,(R)$ form a Witt-complex over R. We use these structures to prove an étale descent for multivariate additive higher Chow groups.

2010 Mathematics Subject Classification: Primary 14C25; Secondary 13F35, 19E15
Keywords and Phrases: algebraic cycle, additive higher Chow group, Witt vectors, de Rham-Witt complex

1. Introduction

The additive higher Chow groups $\mathrm{TCH}^q(X, n; m)$ emerged originally in [5] in part as an attempt to understand certain relative higher algebraic K-groups of schemes in terms of algebraic cycles. Since then, several papers [16], [17], [18], [19], [26], [27], [28] have studied various aspects of these groups. But lack of a suitable moving lemma for smooth *affine* varieties has been a hindrance in studies of their local behaviors. Its projective sibling was known by [17]. During the period of stagnation, the subject has evolved into the notion of 'cycles with modulus' $\mathrm{CH}^q(X|D, n)$ by Binda-Kerz-Saito in [1], [15] associated to pairs (X, D) of schemes and effective Cartier divisors D, setting a more flexible ground, while this desired moving lemma for the affine case was obtained by W. Kai [14] (See Theorem 4.1).

The above developments now propel the authors to continue their program of realizing the relative K-theory $K_n(X \times \mathrm{Spec}\,k[t]/(t^{m+1}), (t))$ in terms of additive higher Chow groups. More specifically, one of the aims in the program considered in this paper is to understand via additive higher Chow groups, the part of the above relative K-groups which was proven in [2] to give the crystalline cohomology. This part turned out to be isomorphic to the de Rham-Witt complexes as seen in [12]. This article is the first of the authors' papers that relate the additive higher Chow groups to the big de Rham-Witt complexes $\mathbb{W}_m\Omega^\bullet_R$ of [8] and to the crystalline cohomology theory. This gives a motivic description of the latter two objects.

While the general notion of cycles with modulus for (X, D) provides a wider picture, the additive higher Chow groups still have a non-trivial operation not shared by the general case. One such is an analogue of the Pontryagin product on homology groups of Lie groups, which turns the additive higher Chow groups into a differential graded algebra (DGA). This product is induced by the structure of algebraic groups on $\mathbb{A}^1$ and $\mathbb{G}_m$ and their action on $X \times \mathbb{A}^r =: X[r]$ for $r \geq 1$.

The usefulness of such a product was already observed in the earliest papers on additive 0-cycles by Bloch-Esnault [5] and Rülling [28]. This product on higher dimensional additive higher Chow cycles was given in [19] for smooth projective varieties. In §5 of this paper, we extend this product structure in two directions: (1) toward multivariate additive higher Chow groups and (2) on smooth affine varieties. In doing so, we generalize some of the necessary tools, such as the following normalization theorem, proven as Theorem 3.2. Necessary definitions are recalled in §2.

Theorem 1.1. *Let X be a smooth scheme which is either quasi-projective or essentially of finite type over a field k. Let D be an effective Cartier divisor on X. Then each cycle class in $\mathrm{CH}^q(X|D, n)$ has a representative, all of whose codimension 1 faces are trivial.*

The above theorem for ordinary higher Chow groups was proven by Bloch and has been a useful tool in dealing with algebraic cycles. In this paper, we use the above theorem to construct the following structure of differential graded algebra and differential graded modules on the multivariate additive higher Chow groups, where Theorem 1.2 is proven in Theorems 7.1, 7.10, and 7.11, while Theorem 1.3 is proven in Theorem 6.13.

Theorem 1.2. *Let X be a smooth scheme which is either affine essentially of finite type or projective over a perfect field k of characteristic $\neq 2$*

(1) *The additive higher Chow groups $\{\mathrm{TCH}^q(X, n; m)\}_{q,n,m \in \mathbb{N}}$ has a functorial structure of a restricted Witt-complex over k.*
(2) *If $X = \mathrm{Spec}\,(R)$ is affine, then $\{\mathrm{TCH}^q(X, n; m)\}_{q,n,m \in \mathbb{N}}$ has a structure of a restricted Witt-complex over R.*
(3) *For X as in (2), there is a natural map of restricted Witt-complexes $\tau^R_{n,m} : \mathbb{W}_m\Omega^{n-1}_R \to \mathrm{TCH}^n(R, n; m)$.*

Theorem 1.3. *Let $r \geq 1$. For a smooth scheme X which is either affine essentially of finite type or projective over a perfect field k of characteristic $\neq 2$, the multivariate additive higher Chow groups $\{\mathrm{CH}^q(X[r]|D_{\underline{m}}, n)\}_{q,n \geq 0}$ with modulus $\underline{m} = (m_1, \cdots, m_r)$, where $m_i \geq 1$, form a differential graded module over* *the DGA* $\{\mathrm{TCH}^q(X, n; |\underline{m}| - 1)\}_{q,n\geq 1}$, *where* $|\underline{m}| = \sum_{i=1}^r m_i$. *In particular, each* $\mathrm{CH}^q(X[r]|D_{\underline{m}}, n)$ *is a* $\mathbb{W}_{(|\underline{m}|-1)}(R)$*-module, when* $X = \mathrm{Spec}\,(R)$ *is affine.*

The above structures on the univariate and multivariate additive higher Chow groups suggest an expectation that these groups may describe the algebraic K-theory relative to nilpotent thickenings of the coordinate axes in an affine space over a smooth scheme. The calculations of such relative K-theory by Hesselholt in [9] and [10] show that any potential motivic cohomology which describes the above relative K-theory may have such a structure.

As part of our program of connecting the additive higher Chow groups with the relative K-theory, we show in [22] that the above map $\tau^R_{n,m}$ is an isomorphism when X is semi-local in addition, and we show how one deduces crystalline cohomology from additive higher Chow groups. The results of this paper form a crucial part in the process.

Recall that the higher Chow groups of Bloch and algebraic K-theory do not satisfy étale descent with integral coefficients. As an application of Theorem 1.3, we show that the étale descent is actually true for the multivariate additive higher Chow groups in the following setting:

Theorem 1.4. *Let* $r \geq 1$ *and let* X *be a smooth scheme which is either affine essentially of finite type or projective over a perfect field* k *of characteristic* $\neq 2$. *Let* G *be a finite group of order prime to* $\mathrm{char}(k)$, *acting freely on* X *with the quotient* $f : X \to X/G$. *Then for all* $q, n \geq 0$ *and and* $\underline{m} = (m_1, \cdots, m_r)$ *with* $m_i \geq 1$ *for* $1 \leq i \leq r$, *the pull-back map* f^* *induces an isomorphism*

$$\mathrm{CH}^q(X/G[r]|D_{\underline{m}}, n) \xrightarrow{\simeq} \mathrm{H}^0(G, \mathrm{CH}^q(X[r]|D_{\underline{m}}, n)).$$

Note that the quotient X/G exists under the hypothesis on X. Since the corresponding descent is not yet known for the relative K-theory of nilpotent thickenings of the coordinate axes in an affine space over a smooth scheme, the above theorem suggests that this descent could be indeed true for the relative K-theory.

Conventions. In this paper, k will denote the base field which will be assumed to be perfect after §4. A k-scheme is a separated scheme of finite type over k. A k-variety is a reduced k-scheme. The product $X \times Y$ means usually $X \times_k Y$, unless said otherwise. We let $\mathbf{Sch}_k$ be the category of k-schemes, $\mathbf{Sm}_k$ of smooth k-schemes, and $\mathbf{SmAff}_k$ of smooth affine k-schemes. A scheme essentially of finite type is a scheme obtained by localizing at a finite subset (including $\emptyset$) of a finite type k-scheme. For $\mathcal{C} = \mathbf{Sch}_k, \mathbf{Sm}_k, \mathbf{SmAff}_k$, we let $\mathcal{C}^{\mathrm{ess}}$ be the extension of the category $\mathcal{C}$ obtained by localizing at a finite subset (including $\emptyset$) of objects in $\mathcal{C}$. We let $\mathbf{SmLoc}_k$ be the category of smooth semi-local k-schemes essentially of finite type over k. So, $\mathbf{SmAff}^{\mathrm{ess}}_k = \mathbf{SmAff}_k \cup \mathbf{SmLoc}_k$ for the objects. When we say a semi-local k-scheme, we always mean one that is essentially of finite type over k. Let $\mathbf{SmProj}_k$ be the category of smooth projective k-schemes.

2. Recollection of basic definitions

For $\mathbb{P}^1 = \mathrm{Proj}_k(k[s_0, s_1])$, we let $y = s_1/s_0$ its coordinate. Let $\square := \mathbb{P}^1 \setminus \{1\}$. For $n \geq 1$, let $(y_1, \cdots, y_n) \in \square^n$ be the coordinates. A face $F \subset \square^n$ means a closed subscheme defined by the set of equations of the form $\{y_{i_1} = \epsilon_1, \cdots, y_{i_s} = \epsilon_s\}$ for an increasing sequence $\{i_j | 1 \leq j \leq s\} \subset \{1, \cdots, n\}$ and $\epsilon_j \in \{0, \infty\}$. We allow $s = 0$, in which case $F = \square^n$. Let $\overline{\square} := \mathbb{P}^1$. A face of $\overline{\square}^n$ is the closure of a face in $\square^n$. For $1 \leq i \leq n$, let $F^1_{n,i} \subset \overline{\square}^n$ be the closed subscheme given by $\{y_i = 1\}$. Let $F^1_n := \sum_{i=1}^n F^1_{n,i}$, which is the cycle associated to the closed subscheme $\overline{\square}^n \setminus \square^n$. Let $\square^0 = \overline{\square}^0 := \mathrm{Spec}\,(k)$. Let $\iota_{n,i,\epsilon} : \square^{n-1} \hookrightarrow \square^n$ be the inclusion $(y_1, \cdots, y_{n-1}) \mapsto (y_1, \cdots, y_{i-1}, \epsilon, y_i, \cdots, y_{n-1})$.

2.1. Cycles with modulus. Let $X \in \mathbf{Sch}^{\mathrm{ess}}_k$. Recall ([21, §2]) that for effective Cartier divisors D_1 and D_2 on X, we say $D_1 \leq D_2$ if $D_1 + D = D_2$ for some effective Cartier divisor D on X. A *scheme with an effective divisor* (sed) is a pair (X, D), where $X \in \mathbf{Sch}^{\mathrm{ess}}_k$ and D an effective Cartier divisor. A morphism $f : (Y, E) \to (X, D)$ of seds is a morphism $f : Y \to X$ in $\mathbf{Sch}^{\mathrm{ess}}_k$ such that $f^*(D)$ is defined as a Cartier divisor on Y and $f^*(D) \leq E$. In particular, $f^{-1}(D) \subset E$. If $f : Y \to X$ is a morphism of k-schemes, and (X, D) is a sed such that $f^{-1}(D) = \emptyset$, then $f : (Y, \emptyset) \to (X, D)$ is a morphism of seds.

Definition 2.1 ([1], [15]). Let (X, D) and $(\overline{Y}, E)$ be schemes with effective divisors. Let $Y = \overline{Y} \setminus E$. Let $V \subset X \times Y$ be an integral closed subscheme with closure $\overline{V} \subset X \times \overline{Y}$. We say V *has modulus D (relative to E)* if $\nu^*_V(D \times \overline{Y}) \leq \nu^*_V(X \times E)$ on $\overline{V}^N$, where $\nu_V : \overline{V}^N \to \overline{V} \hookrightarrow X \times \overline{Y}$ is the normalization followed by the closed immersion.

Recall the following containment lemma from [21, Proposition 2.4] (see also [1, Lemma 2.1] and [17, Proposition 2.4]):

Proposition 2.2. *Let (X, D) and $(\overline{Y}, E)$ be schemes with effective divisors and $Y = \overline{Y} \setminus E$. If $V \subset X \times Y$ is a closed subscheme with modulus D relative to E, then any closed subscheme $W \subset V$ also has modulus D relative to E.*

Definition 2.3 ([1], [15]). Let (X, D) be a scheme with an effective divisor. For $s \in \mathbb{Z}$ and $n \geq 0$, let $\underline{z}_s(X|D, n)$ be the free abelian group on integral closed subschemes $V \subset X \times \square^n$ of dimension $s + n$ satisfying the following conditions:

(1) (Face condition) for each face $F \subset \square^n$, V intersects $X \times F$ properly.
(2) (Modulus condition) V has modulus D relative to F^1_n on $X \times \square^n$.

We usually drop the phrase "relative to F^1_n" for simplicity. A cycle in $\underline{z}_s(X|D, n)$ is called an *admissible cycle with modulus D*. One checks that $(n \mapsto \underline{z}_s(X|D, n))$ is a cubical abelian group. In particular, the groups $\underline{z}_s(X|D, n)$ form a complex with the boundary map $\partial = \sum_{i=1}^n (-1)^i (\partial^\infty_i - \partial^0_i)$, where $\partial^\epsilon_i = \iota^*_{n,i,\epsilon}$.

Definition 2.4 ([1], [15]). The complex $(z_s(X|D, \bullet), \partial)$ is the nondegenerate complex associated to $(n \mapsto \underline{z}_s(X|D, n))$, i.e., $z_s(X|D, n) :=$

$\underline{z}_s(X|D,n)/\underline{z}_s(X|D,n)_{\text{degn}}$. The homology $\mathrm{CH}_s(X|D,n) := \mathrm{H}_n(z_s(X|D,\bullet))$ for $n \geq 0$ is called *higher Chow group of X with modulus D*. If X is equidimensional of dimension d, for $q \geq 0$, we write $\mathrm{CH}^q(X|D,n) = \mathrm{CH}_{d-q}(X|D,n)$.

Here is a special case from [21]:

Definition 2.5. Let $X \in \mathbf{Sch}_k^{\mathrm{ess}}$. For $r \geq 1$, let $X[r] := X \times \mathbb{A}^r$. When $(t_1, \cdots, t_r) \in \mathbb{A}^r$ are the coordinates, and $m_1, \cdots, m_r \geq 1$ are integers, let $D_{\underline{m}}$ be the divisor on $X[r]$ given by the equation $\{t_1^{m_1} \cdots t_r^{m_r} = 0\}$. The groups $\mathrm{CH}^q(X[r]|D_{\underline{m}}, n)$ are called *multivariate additive higher Chow groups* of X. For simplicity, we often say "a cycle with modulus $\underline{m}$" for "a cycle with modulus $D_{\underline{m}}$." For an r-tuple of integers $\underline{m} = (m_1, \cdots, m_r)$, we write $|\underline{m}| = \sum_{i=1}^r m_i$. We shall say that $\underline{m} \geq p$ if $m_i \geq p$ for each i.
When $r = 1$, we obtain additive higher Chow groups, and as in [19], we often use the older notations $\mathrm{Tz}^q(X, n+1; m-1)$ for $z^q(X[1]|D_m, n)$ and $\mathrm{TCH}^q(X, n+1; m-1)$ for $\mathrm{CH}^q(X[1]|D_m, n)$. In such cases, note that the modulus m is shifted by 1 from the above sense.

Definition 2.6. Let $\mathcal{W}$ be a finite set of locally closed subsets of X and let $e : \mathcal{W} \to \mathbb{Z}_{\geq 0}$ be a set function. Let $\underline{z}^q_{\mathcal{W},e}(X|D,n)$ be the subgroup generated by integral cycles $Z \in \underline{z}^q(X|D,n)$ such that for each $W \in \mathcal{W}$ and each face $F \subset \square^n$, we have $\mathrm{codim}_{W \times F}(Z \cap (W \times F)) \geq q - e(W)$. They form a subcomplex $\underline{z}^q_{\mathcal{W},e}(X|D,\bullet)$ of $\underline{z}^q(X|D,\bullet)$. Modding out by degenerate cycles, we obtain the subcomplex $z^q_{\mathcal{W},e}(X|D,\bullet) \subset z^q(X|D,\bullet)$. We write $z^q_{\mathcal{W}}(X|D,\bullet) := z^q_{\mathcal{W},0}(X|D,\bullet)$. For additive higher Chow cycles, we write $\underline{\mathrm{Tz}}^q_{\mathcal{W}}(X,n;m)$ for $\underline{z}^q_{\mathcal{W}[1]}(X[1]|D_{m+1}, n-1)$, where $\mathcal{W}[1] = \{W[1] \mid W \in \mathcal{W}\}$.

Here are some basic lemmas used in the paper:

Lemma 2.7 ([21, Lemma 2.2]). *Let $f : Y \to X$ be a dominant map of normal integral k-schemes. Let D be a Cartier divisor on X such that the generic points of* $\mathrm{Supp}(D)$ *are contained in $f(Y)$. Suppose that $f^*(D) \geq 0$ on Y. Then $D \geq 0$ on X.*

Lemma 2.8 ([21, Lemma 2.9]). *Let $f : Y \to X$ be a proper morphism of quasi-projective k-varieties. Let $D \subset X$ be an effective Cartier divisor such that $f(Y) \not\subset D$. Let $Z \in z^q(Y|f^*(D), n)$ be an irreducible cycle. Let $W = f(Z)$ on $X \times \square^n$. Then $W \in z^s(X|D,n)$, where $s = \mathrm{codim}_{X \times \square^n}(W)$.*

Lemma 2.9. *Let X be a k-scheme, and let $\{U_i\}_{i \in I}$ be an open cover of X. Let $Z \in z^q(X \times \square^n)$ and let Z_{U_i} be the flat pull-back to $U_i \times \square^n$. Then $Z \in z^q(X|D,n)$ if and only if for each $i \in I$, we have $Z_{U_i} \in z^q(U_i|D_{U_i}, n)$, where D_{U_i} is the restriction of D on U_i.*

Proof. The direction ($\Rightarrow$) is obvious since flat pull-backs respect admissibility of cycles with modulus by [21, Proposition 2.12]. For the direction ($\Leftarrow$), we may assume Z is irreducible. In this case, it is easily checked that the face and the modulus conditions are both local on the base X. $\square$

2.2. de Rham-Witt complexes.

2.2.1. *Ring of big Witt-vectors.* Let R be a commutative ring with unit. We recall the definition of the ring of big Witt-vectors of R (see [11, §4] or [28, Appendix A]). A *truncation set* $S \subset \mathbb{N}$ is a non-empty subset such that if $s \in S$ and $t|s$, then $t \in S$. As a set, let $\mathbb{W}_S(R) := R^S$ and define the map $w : \mathbb{W}_S(R) \to R^S$ by sending $a = (a_s)_{s\in S}$ to $w(a) = (w(a)_s)_{s\in S}$, where $w(a)_s := \sum_{t|s} t a_t^{s/t}$. When R^S on the target of w is given the component-wise ring structure, it is known that there is a unique functorial ring structure on $\mathbb{W}_S(R)$ such that w is a ring homomorphism (see [11, Proposition 1.2]). When $S = \{1, \cdots, m\}$, we write $\mathbb{W}_m(R) := \mathbb{W}_S(R)$.
There is another description. Let $\mathbb{W}(R) := \mathbb{W}_{\mathbb{N}}(R)$. Consider the multiplicative group $(1 + tR[[t]])^\times$, where t is an indeterminate. Then there is a natural bijection $\mathbb{W}(R) \simeq (1+tR[[t]])^\times$, where the addition in $\mathbb{W}(R)$ corresponds to the multiplication of formal power series. For a truncation set S, we can describe $\mathbb{W}_S(R)$ as the quotient of $(1+tR[[t]])^\times$ by a suitable subgroup I_S. See [28, A.7] for details. In case $S = \{1, \cdots, m\}$, we can write $\mathbb{W}_m(R) = (1+tR[[t]])^\times/(1+t^{m+1}R[[t]])^\times$ as an additive group.
For $a \in R$, the Teichmüller lift $[a] \in \mathbb{W}_S(R)$ corresponds to the image of $1 - at \in (1 + tR[[t]])^\times$. This yields a multiplicative map $[-] : R \to \mathbb{W}_S(R)$. The additive identity element of $\mathbb{W}_m(R)$ corresponds to the unit polynomial 1 and the multiplicative identity element corresponds to the polynomial $1 - t$.

2.2.2. *de Rham-Witt complex.* Let p be an odd prime and R be a $\mathbb{Z}_{(p)}$-algebra.[1] For each truncation set S, there is a differential graded algebra $\mathbb{W}_S\Omega^\bullet_R$ called the big de Rham-Witt complex over R. This defines a contravariant functor on the category of truncation sets. This is an initial object in the category of V-complexes and in the category of Witt-complexes over R. For details, see [8] and [28, §1]. When S is a finite truncation set, we have $\mathbb{W}_S\Omega^\bullet_R = \Omega^\bullet_{\mathbb{W}_S(R)/\mathbb{Z}}/N^\bullet_S$, where $N^\bullet_S$ is the differential graded ideal given by some generators ([28, Proposition 1.2]). In case $S = \{1, 2, \cdots, m\}$, we write $\mathbb{W}_m\Omega^\bullet_R$ for this object.
Here is another relevant object for this paper from [8, Definition 1.1.1]; *a restricted Witt-complex over R* is a pro-system of differential graded $\mathbb{Z}$-algebras $((E_m)_{m\in\mathbb{N}}, \mathfrak{R} : E_{m+1} \to E_m)$, with homomorphisms of graded rings $(F_r : E_{rm+r-1} \to E_m)_{m,r\in\mathbb{N}}$ called the *Frobenius* maps, and homomorphisms of graded groups $(V_r : E_m \to E_{rm+r-1})_{m,r\in\mathbb{N}}$ called the *Verschiebung* maps, satisfying the following relations for all $n, r, s \in \mathbb{N}$:

(i) $\mathfrak{R}F_r = F_r\mathfrak{R}^r, \mathfrak{R}^rV_r = V_r\mathfrak{R}, F_1 = V_1 = \mathrm{Id}, F_rF_s = F_{rs}, V_rV_s = V_{rs}$;
(ii) $F_rV_r = r$. When $(r, s) = 1$, $F_rV_s = V_sF_r$ on E_{rm+r-1};
(iii) $V_r(F_r(x)y) = xV_r(y)$ for all $x \in E_{rm+r-1}$ and $y \in E_m$; (projection formula)
(iv) $F_rdV_r = d$, where d is the differential of the DGAs.

Furthermore, we require that there is a homomorphism of pro-rings $(\lambda : \mathbb{W}_m(R) \to E^0_m)_{m\in\mathbb{N}}$ that commutes with F_r and V_r, satisfying

[1] A definition of Witt-complex over a more general ring R can be found in [11, Definition 4.1].

(v) $F_r d\lambda([a]) = \lambda([a]^{r-1}) d\lambda([a])$ for all $a \in R$ and $r \in \mathbb{N}$.

The pro-system $\{\mathbb{W}_m \Omega^\bullet_R\}_{m \geq 1}$ is the initial object in the category of restricted Witt-complexes over R (See [28, Proposition 1.15]).

3. Normalization theorem

Let k be any field. The aim of this section is to prove Theorem 3.2. Such results were known when $D = \emptyset$, or when X is replaced by $X \times \mathbb{A}^1$ with $D = \{t^{m+1} = 0\}$ for $t \in \mathbb{A}^1$. We generalize it to higher Chow groups with modulus.

Definition 3.1. Let (X, D) be a scheme with an effective divisor. Let $z^q_N(X|D, n)$ be the subgroup of cycles $\alpha \in z^q(X|D, n)$ such that $\partial^0_i(\alpha) = 0$ for all $1 \leq i \leq n$ and $\partial^\infty_i(\alpha) = 0$ for $2 \leq i \leq n$. One checks that $\partial^\infty_1 \circ \partial^\infty_1 = 0$. Writing ∂^∞_1 as ∂^N, we obtain a subcomplex $\iota : (z^q_N(X|D, \bullet), \partial^N) \hookrightarrow (z^q(X|D, \bullet), \partial)$.

Theorem 3.2. *Let $X \in \mathbf{Sm}^{\mathrm{ess}}_k$ and let $D \subset X$ be an effective Cartier divisor. Then $\iota : z^q_N(X|D, \bullet) \to z^q(X|D, \bullet)$ is a quasi-isomorphism. In particular, every cycle class in $\mathrm{CH}^q(X|D, n)$ can be represented by a cycle α such that $\partial^\epsilon_i(\alpha) = 0$ for all $1 \leq i \leq n$ and $\epsilon = 0, \infty$.*

Let Cube be the standard category of cubes (see [24, §1]) so that a cubical abelian group is a functor $\mathrm{Cube}^{\mathrm{op}} \to (\mathbf{Ab})$. Recall also from *loc.cit.* that an extended cubical abelian is a functor $\mathrm{ECube}^{\mathrm{op}} \to (\mathbf{Ab})$, where ECube is the smallest symmetric monoidal subcategory of Sets containing Cube and the morphism $\mu : \underline{2} \to \underline{1}$. The essential point of the proof of Theorem 3.2 is

Theorem 3.3. *Let $X \in \mathbf{Sm}^{\mathrm{ess}}_k$ and $D \subset X$ be an effective Cartier divisor. Then $(\underline{n} \mapsto z^q(X|D; n))$ is an extended cubical abelian group.*

If Theorem 3.3 holds, then [24, Lemma 1.6] implies Theorem 3.2. We suppose (X, D) is as in Theorem 3.2 in what follows. The idea is similar to that of [19, Appendix].

Let $q_1 : \square^2 \to \square$ be the morphism $(y_1, y_2) \mapsto y_1 + y_2 - y_1 y_2$ if $y_1, y_2 \neq \infty$, and $(y_1, y_2) \mapsto \infty$ if y_1 or $y_2 = \infty$. Under the identification $\psi : \square \simeq \mathbb{A}^1$ given by $y \mapsto 1/(1-y)$ (which sends $\{\infty, 0\}$ to $\{0, 1\}$), this map q_1 is equivalent to $q_{1,\psi} : \mathbb{A}^2 \to \mathbb{A}^1$ given by $(y_1, y_2) \mapsto y_1 y_2$. For our convenience, we use this $\square_\psi := (\mathbb{A}^1, \{0, 1\})$ and cycles on $X \times \square^n_\psi$. The boundary operator is $\partial = \sum_{i=1}^n (-1)^i (\partial^0_i - \partial^1_i)$, and we replace $F^1_{n,i}$ by $F^\infty_{n,i} = \{y_i = \infty\}$. We write $F^\infty_n = \sum_{i=1}^n F^\infty_{n,i}$. We write $\overline{\square}_\psi = (\mathbb{P}^1, \{0, 1\})$. The group of admissible cycles is $\underline{z}^q_\psi(X|D, n)$. Consider $q_{n,\psi} : X \times \square^{n+1}_\psi \to X \times \square^n_\psi$ given by $(x, y_1, \cdots, y_{n+1}) \mapsto (x, y_1, \cdots, y_{n-1}, y_n y_{n+1})$.

Proposition 3.4. *For $Z \in z^q_\psi(X|D, n)$, we have $q^*_{n,\psi}(Z) \in z^q_\psi(X|D, n+1)$.*

The delicacy of its proof lies in that the product map $q_{1,\psi} : \mathbb{A}^2 \to \mathbb{A}^1$ *does not* extend to a morphism $(\mathbb{P}^1)^2 \to \mathbb{P}^1$ of varieties so that checking the modulus condition becomes nontrivial. We use a correspondence instead. For $n \geq 1$, let

$i_n : W_n \hookrightarrow X \times \square_{\psi}^{n+1} \times \overline{\square}_{\psi}^{1}$ be the closed subscheme defined by the equation $u_0 y_n y_{n+1} = u_1$, where $(y_1, \cdots, y_{n+1}) \in \square_{\psi}^{n+1}$ and $(u_0; u_1) \in \overline{\square}_{\psi}^{1}$ are the coordinates. Let $y := u_1/u_0$. Its Zariski closure $\overline{W}_n \hookrightarrow X \times \overline{\square}_{\psi}^{n+1} \times \overline{\square}_{\psi}^{1}$ is given by the equation $u_0 u_{n,1} u_{n+1,1} = u_1 u_{n,0} u_{n+1,0}$, where $(u_{1,0}, u_{1,1}), \cdots, (u_{n+1,0}, u_{n+1,1})$ are the homogeneous coordinates of $\overline{\square}_{\psi}^{n+1}$ with $y_i = u_{i,1}/u_{i,0}$.

Consider $\theta_n : X \times \square_{\psi}^{n+1} \times \overline{\square}_{\psi}^{1} \to X \times \square_{\psi}^{n}$ given by $(x, y_1, \cdots, y_{n+1}, (u_0; u_1)) \mapsto (x, y_1, \cdots, y_{n-1}, y_n y_{n+1})$, and let $\pi_n := \theta_n|_{W_n}$. To extend this π_n to a morphism $\overline{\pi}_n$ on $\overline{W}_n$, we use the projection $\overline{\theta}_n : X \times \overline{\square}_{\psi}^{n+1} \times \overline{\square}_{\psi}^{1} \to X \times \overline{\square}_{\psi}^{n-1} \times \overline{\square}_{\psi}^{1}$, that drops the coordinates $(u_{n,0}; u_{n,1})$ and $(u_{n+1,0}; u_{n+1,1})$, and the projection $p_n : X \times \square_{\psi}^{n+1} \times \overline{\square}_{\psi}^{1} \to X \times \square_{\psi}^{n+1}$, that drops the last coordinate $(u_0; u_1)$.

Lemma 3.5. (1) *$W_n \cap \{u_0 = 0\} = \emptyset$, so that $W_n \subset X \times \square_{\psi}^{n+1} \times \square_{\psi}^{1}$.* (2) *$\overline{\theta}_n|_{W_n} = \pi_n$. Thus, we define $\overline{\pi}_n := \overline{\theta}_n|_{\overline{W}_n}$, which extends π_n.* (3) *The varieties W_n and $\overline{W}_n$ are smooth.* (4) *Both π_n and $\overline{\pi}_n$ are surjective flat morphisms of relative dimension* 1.

Proof. Its proof is almost identical to that of [19, Lemma A.5]. Part (1) follows from the defining equation of W_n, and (2) holds by definition. Let $\rho_n := p_n|_{W_n} : W_n \to X \times \square_{\psi}^{n+1}$. Since X is smooth, using Jacobian criterion we check that W_n is smooth. Furthermore, ρ_n is an isomorphism with the obvious inverse. Under this identification, the morphism π_n can also be regarded as the projection $(x, y_1, \cdots, y_n, y) \mapsto (x, y_1, \cdots, y_{n-1}, y)$ that drops y_n. In particular, π_n is a smooth and surjective of relative dimension 1. To check that $\overline{W}_n$ is smooth, one can do it locally on each open set where each of $u_{n,i}, u_{n+1,i}, u_i$ is nonzero for $i = 0, 1$. In each such open set, the equation for $\overline{W}_n$ takes the same form as for W_n, so that it is smooth again by Jacobian criterion. Similarly as for π_n, one sees $\overline{\pi}_n$ is of relative dimension 1. Since $\overline{\theta}_n$ is projective and π_n is surjective, the morphism $\overline{\pi}_n$ is projective and surjective. So, since $\overline{W}_n$ is smooth, the map $\overline{\pi}_n$ is flat by [7, Exercise III-10.9, p.276]. Thus, we have (3) and (4). $\square$

Lemma 3.6. *Let $n \geq 1$ and let $Z \subset X \times \square_{\psi}^{n}$ be a closed subscheme with modulus D. Then $Z' := (i_n)_*(\pi_n^*(Z))$ also has modulus D.*

Proof. Let $\overline{Z}$ and $\overline{Z}'$ be the Zariski closures of Z and Z' in $X \times \overline{\square}_{\psi}^{n}$ and $X \times \overline{\square}_{\psi}^{n+1}$, respectively. By Lemma 3.5 and the projectivity of $\overline{\theta}_n$, we see that $\overline{\theta}_n(\overline{Z}') = \overline{Z}$. Consider the commutative diagram

$$\begin{tikzcd}
\overline{Z}'^N \arrow[r, "g"] \arrow[d, "f"] \arrow[rr, bend left, "\nu_{Z'}"] & \overline{W}_n \arrow[r, hook, "\overline{i}_n"] \arrow[dr, "\overline{\pi}_n"] & X \times \overline{\square}_{\psi}^{n+1} \times \overline{\square}_{\psi}^{1} \arrow[d, "\overline{\theta}_n"] \\
\overline{Z}^N \arrow[rr, "\nu_Z"] & & X \times \overline{\square}_{\psi}^{n},
\end{tikzcd} \tag{3.1}$$

where f is induced by the surjection $\overline{\theta}_n|_{\overline{Z}'} : \overline{Z}' \to \overline{Z}$, the maps g and ν_Z are normalizations of $\overline{Z}'$ and $\overline{Z}$ composed with the closed immersions, and $\nu_{Z'} := \overline{i}_n \circ g$. By the definition of $\overline{\theta}_n$, we have $\overline{\theta}_n^*(D \times \overline{\square}_\psi^n) = D \times \overline{\square}_\psi^{n+2}$, $\overline{\theta}_n^*(F_{n,n}^\infty) = F_{n+2,n+2}^\infty$, while $\overline{\theta}_n^*(F_{n,i}^\infty) = F_{n+2,i}^\infty$ for $1 \leq i \leq n-1$. By the defining equation of $\overline{W}_n$, we have $\overline{\pi}_n^* F_{n,n}^\infty = \overline{i}_n^* F_{n+2,n+2}^\infty = \overline{i}_n^*\{u_0 = 0\} \leq \overline{i}_n^*(\{u_{n,0} = 0\} + \{u_{n+1,0} = 0\}) = \overline{i}_n^*(F_{n+2,n}^\infty + F_{n+2,n+1}^\infty)$.
Thus, $\nu_{Z'}^* \overline{\theta}_n^* \sum_{i=1}^n F_{n,i}^\infty = \sum_{i=1}^{n-1} \nu_{Z'}^* F_{n+2,i}^\infty + g^* \overline{\pi}_n^* F_{n,n}^\infty \leq \sum_{i=1}^{n-1} \nu_{Z'}^* F_{n+2,i}^\infty + g^* \overline{i}_n^*(F_{n+2,n}^\infty + F_{n+2,n+1}^\infty) = \sum_{i=1}^{n+1} \nu_{Z'}^* F_{n+2,i}^\infty \leq \sum_{i=1}^{n+2} \nu_{Z'}^* F_{n+2,i}^\infty$. (In case $n=1$, we just ignore the terms with $\sum_{i=1}^{n-1}$ in the above.)
That Z has modulus D means $\nu_Z^*(D \times \overline{\square}_\psi^n) \leq \sum_{i=1}^n \nu_Z^* F_{n,i}^\infty$. Applying f^* and using (3.1), we have $\nu_{Z'}^*(D \times \overline{\square}_\psi^{n+2}) = \nu_{Z'}^* \overline{\theta}_n^*(D \times \overline{\square}_\psi^n) \leq \nu_{Z'}^* \overline{\theta}_n^* \sum_{i=1}^n F_{n,i}^\infty$, which is bounded by $\sum_{i=1}^{n+2} \nu_{Z'}^* F_{n+2,i}^\infty$ as we saw above. This means Z' has modulus D. □

Definition 3.7. For any closed subscheme $Z \subset X \times \square_\psi^n$, we define $W_n(Z) := p_{n*} i_{n*} \pi_n^*(Z)$, which is closed in $X \times \square_\psi^{n+1}$.

Lemma 3.8. *Let $n \geq 1$. If a closed subscheme $Z \subset X \times \square_\psi^n$ intersects all faces properly, then $W_n(Z)$ intersects all faces of $X \times \square_\psi^{n+1}$ properly.*

Proof. Our W_n is equal to $\tau^* \tau_n^* \tau_{n+1}^* W_n^X$, where W_n^X is that of [23, Lemma 4.1], and τ, τ_n, τ_{n+1} are the involutions $(x \mapsto 1-x)$ for y, y_n, y_{n+1}, respectively. So, the lemma is a special case of *loc.cit.* □

Proof of Proposition 3.4. Consider the commutative diagram

$$\begin{tikzcd}
W_n \arrow[r, hook, "i_n"] \arrow[d, "\pi_n"'] \arrow[dr, "\rho_n = p_n|_{W_n}"] & X \times \square_\psi^{n+1} \times \overline{\square}_\psi \arrow[d, "p_n"] \\
X \times \square_\psi^n & X \times \square_\psi^{n+1}. \arrow[l, "q_{n,\psi}"']
\end{tikzcd}$$

By Lemma 3.5, ρ_n is an isomorphism so that $\rho_{n*} i_n^* p_n^* = \mathrm{Id}$. Hence, $q_{n,\psi}^*(Z) = \rho_{n*} i_n^* p_n^* q_{n,\psi}^*(Z) =^\dagger \rho_{n*} \pi_n^*(Z) =^\ddagger p_{n*} i_{n*} \pi_n^*(Z) = W_n(Z)$, where †, ‡ are due to commutativity. So, we have reduced to showing that $W_n(Z) \in z_\psi^q(X|D, n+1)$. But, by Lemmas 3.6 and 3.8, we have $i_{n*}\pi_n^*(Z) \in z_\psi^{q+1}(X \times \mathbb{P}^1 | D \times \mathbb{P}^1, n+1)$. Now, for the projection p_n, by Lemma 2.8, we have $W_n(Z) = p_{n*} i_{n*} \pi_n^*(Z) \in z_\psi^q(X|D, n+1)$. This proves Proposition 3.4. □

Proof of Theorem 3.3. Since we know that $(\underline{n} \mapsto z^q(X|D; n))$ is a cubical abelian group, every morphism $h : \underline{r} \to \underline{s}$ in Cube induces a morphism $h : \square^r \to \square^s$ which gives a homomorphism $h^* : z^q(X|D, s) \to z^q(X|D, r)$. Furthermore, the morphism $\mu : \underline{2} \to \underline{1}$ induces the morphism $q_1 : \square^2 \to \square^1$ of varieties, and for each $Z \in z^q(X|D, 1)$, we have $q_1^*(Z) \in z^q(X|D, 2)$. Indeed, under the isomorphism $\psi : \square \simeq \mathbb{A}^1, y \mapsto 1/(1-y)$, this is equivalent to show that $q_{1,\psi}^*$ sends admissible cycles to admissible cycles, which we know by Proposition 3.4.

So, it only remains to show the following "stability under products": if $h_i : \underline{r_i} \to \underline{s_i}$, $i = 1, 2$, are morphisms in ECube such that the corresponding morphisms $h_i : \square^{r_i} \to \square^{s_i}$ induce homomorphisms $h_i^* : z^q(X|D, s_i) \to z^q(X|D, r_i)$, for $i = 1, 2$ and all $q \geq 0$, then $h := h_1 \times h_2 : \square^{r_1+r_2} \to \square^{s_1+s_2}$ induces a homomorphism $h^* : z^q(X|D, s) \to z^q(X|D, r)$ for all $q \geq 0$, where $r = r_1 + r_2$ and $s = s_1 + s_2$.

Since $h = h_1 \times h_2 = (\mathrm{Id}_{r_1} \times h_2) \circ (h_1 \times \mathrm{Id}_{r_2})$, we reduce to prove it when h is either $\mathrm{Id}_{r_1} \times h_2$ or $h_1 \times \mathrm{Id}_{r_2}$. But the statement obviously holds for these cases. $\square$

4. On moving lemmas

Let k be any field. In this section, we discuss some of moving lemmas on algebraic cycles with modulus conditions. By a 'moving lemma', we ask whether the inclusion $z^q_{\mathcal{W}}(Y|D, \bullet) \subset z^q(Y|D, \bullet)$ in Definition 2.6 is a quasi-isomorphism. It is known when Y is smooth quasi-projective and $D = 0$ (by [4]), and when $Y = X \times \mathbb{A}^1$, with X smooth projective, $D = X \times \{t^{m+1} = 0\}$, and $\mathcal{W}$ consists of $W \times \mathbb{A}^1$ for finitely many locally closed subsets $W \subset X$ (by [17]). Recently, W. Kai [14] proved it when Y is smooth affine with a suitable condition. Kai's cases include the above case of $Y = X \times \mathbb{A}^1$, where X is this time smooth affine. His proof applies to more general cases, possibly after Nisnevich sheafifications. In §4.1, we sketch the argument of Kai in the case of multivariate additive higher Chow groups of smooth affine k-variety. In §4.2, we generalize the moving lemma of [17] in the case of pairs $(X \times S, X \times D)$ where X is smooth projective. In §4.3 and 4.4, we discuss the standard pull-back property and its consequences. In §4.5, we discuss a moving lemma for additive higher Chow groups of smooth semi-local k-schemes essentially of finite type.

4.1. Kai's affine method for multivariate additive higher Chow groups.

The moving lemma of W. Kai [14] is the first moving result that applies to cycle groups with a *non-zero* modulus over a smooth *affine* scheme. Since the work *loc. cit.* is at present not yet refereed, we give a detailed sketch the proof of the following special case on multivariate additive higher Chow groups. But, we emphasize that the most crucial part is due to Kai. Following Definition 2.5, we write $X[r] := X \times \mathbb{A}^r$.

Theorem 4.1 (W. Kai). *Let X be a smooth affine variety over any field k. Let $\mathcal{W}$ be a finite set of locally closed subsets of X. Let $\mathcal{W}[r] := \{W[r] \mid W \in \mathcal{W}\}$. Let $\underline{m} = (m_1, \cdots, m_r) \geq 1$. Then the inclusion $z^q_{\mathcal{W}[r]}(X[r]|D_{\underline{m}}, \bullet) \hookrightarrow z^q(X[r]|D_{\underline{m}}, \bullet)$ is a quasi-isomorphism.*

First recall some preparatory results:

Lemma 4.2 ([17, Lemma 4.5]). *Let $f : X \to Y$ be a dominant morphism of normal varieties. Suppose that Y is integral with the generic point $\eta \in Y$, and let X_η be the fiber over η, with the inclusion $j_\eta : X_\eta \hookrightarrow X$. Let D be a Weil* *divisor on X such that $j_\eta^*(D) \geq 0$. Then there exists a non-empty open subset $U \subset Y$ such that $j_U^*(D) \geq 0$, where $j_U : f^{-1}(U) \hookrightarrow X$ is the inclusion.*

The following generalizes [17, Proposition 4.7]:

Proposition 4.3 (Spreading lemma). *Let $k \subset K$ be a purely transcendental extension. Let (X, D) be a smooth quasi-projective k-scheme with an effective Cartier divisor, and let $\mathcal{W}$ be a finite collection of locally closed subsets of X. Let (X_K, D_K) and $\mathcal{W}_K$ be the base changes via* $\mathrm{Spec}\,(K) \to \mathrm{Spec}\,(k)$. *Let $p_{K/k} : X_K \to X_k$ be the base change map. Then the pull-back map*

$$p_{K/k}^* : \frac{z^q(X|D, \bullet)}{z^q_{\mathcal{W}}(X|D, \bullet)} \to \frac{z^q(X_K|D_K, \bullet)}{z^q_{\mathcal{W}_K}(X_K|D_K, \bullet)}$$

is injective on homology.

Proof. It is similar to [17, Proposition 4.7]. We sketch its proof for the reader's convenience. If k is finite, then we can use the standard pro-ℓ-extension argument to reduce the proof to the case when k is infinite, which we assume from now. We may also assume that $\mathrm{tr.deg}_k K < \infty$ and furthermore that $\mathrm{tr.deg}_k K = 1$, by induction. So, we have $K = k(\mathbb{A}^1_k)$.

Suppose $Z \in z^q(X|D, n)$ is a cycle that satisfies $\partial Z \in z^q_{\mathcal{W}}(X|D, n-1)$, and $Z_K = \partial(B_K) + V_K$ for some $B_K \in z^q(X_K|D_K, n+1)$ and $V_K \in z^q_{\mathcal{W}_K}(X_K|D_K, n)$. Consider the inclusion $z^q(X_K|D_K, \bullet) \hookrightarrow z^q(X_K, \bullet)$. Then there is a non-empty open $U' \subset \mathbb{A}^1_k$ such that $B_K = B_{U'}|_\eta, V_K = V_{U'}|_\eta$, $Z \times U' = \partial(B_{U'}) + V_{U'}$ for some $B_{U'} \in z^q(X \times U', n+1)$, $V_{U'} \in z^q_{\mathcal{W} \times U'}(X \times U', n)$, where η is the generic point of U'. Let $j_\eta : X \times \eta \to X \times U'$ be the inclusion, which is flat.

Since B_K, V_K satisfy the modulus condition, we have $j_\eta^*(X \times U' \times F^1_{n+1} - D \times U' \times \overline{\square}^{n+1}) \geq 0$ on $\overline{B}^N_K$ and similarly for $\overline{V}^N_K$. Furthermore, $\overline{B}^N_{U'} \to U', \overline{V}^N_{U'} \to U'$ are dominant. Thus by Lemma 4.2, there is a non-empty open $U \subset U'$ such that $j_U^*(X \times U' \times F^1_{n+1} - D \times U' \times \overline{\square}^{n+1}) \geq 0$ on $\overline{B}^N_U$ and similarly for $\overline{V}^N_U$, for $j_U : X \times U \hookrightarrow X \times U'$. This proves that B_U and V_U have modulus $D \times U$. Hence, $B_U \in z^q(X \times U|D \times U, n+1)$ and $V_U \in z^q_{\mathcal{W} \times U}(X \times U|D \times U, n)$ with $Z \times U = \partial(B_U) + V_U$.

Since k is infinite, the set $U(k) \hookrightarrow U$ is dense. We claim the following:

Claim: *There is a point $u \in U(k)$ such that the pull-backs of B_U and V_U under the inclusion $i_u : X \times \{u\} \hookrightarrow X \times U$ are both defined in $z^q(X, n+1)$ and $z^q_{\mathcal{W}}(X, n)$, respectively.*

Its proof requires the following elementary fact:

Lemma: *Let Y be any k-scheme. Let $B \in z^q(Y \times U)$ be a cycle. Then there exists a nonempty open subset $U'' \subset U$ such that for each $u \in U''(k)$, the closed subscheme $Y \times \{u\}$ intersects B properly on $Y \times U$, thus it defines a cycle $i_u^*(B) \in z^q(Y)$, where Y is identified with $Y \times \{u\}$.*

Note that for each $u \in U(k)$, the subscheme $Y \times \{u\} \subset Y \times U$ is an effective divisor, so its proper intersection with B is equivalent to that $Y \times \{u\}$ does not contain any irreducible component of B. If there exists a point $u_i \in U(k)$

such that $Y \times \{u_i\}$ contains an irreducible component B_i of B, then for any other $u \in U(k) \setminus \{u_i\}$, we have $(Y \times \{u\}) \cap B_i = \emptyset$. So, for every irreducible component B_i of B, there exists at most one $u_i \in U(k)$ such that $Y \times \{u_i\}$ contains B_i. Let S be the union of such points u_i, if they exist. There are only finitely many irreducible components of B, so $|S| < \infty$. Taking $U'' := U \setminus S$, we have Lemma.

We now prove Claim. Let $F \subset \square^{n+1}$ be any face, including the case $F = \square^{n+1}$. Since $B_U \in z^q(X \times U, n+1)$, by definition $X \times U \times F$ and B_U intersect properly on $X \times U \times \square^{n+1}$, so their intersection gives a cycle $B_{U,F} \in z^q(X \times U \times F)$. By Lemma with $Y = X \times F$, there exists a nonempty open subset $U_F \subset U$ such that $B_{U,F}$ defines a cycle in $z^q(X \times \{u\} \times F)$ for every $u \in U_F(k)$. Let $\mathcal{U}_1 := \bigcap_F U_F$, where the intersection is taken over all faces F of $\square^{n+1}$. This is a nonempty open subset of U. Similarly, let $F \subset \square^n$ be any face, including the case $F = \square^n$. Here, $V_U \in z^q_{\mathcal{W} \times U}(X \times U, n)$, and repeating the above argument involving Lemma with $Y = W \times F$ for $W \in \mathcal{W}$, we get a nonempty open subset $U_{W,F} \subset U$ such that we have an induced cycle in $z^q(W \times \{u\} \times F)$ for every $u \in U_{W,F}(k)$. Let $\mathcal{U}_2 := \bigcap_{W,F} U_{W,F}$, where the intersection is taken over all pairs (W, F), with $W \in \mathcal{W}$ and a face $F \subset \square^n$. Taking $\mathcal{U} := \mathcal{U}_1 \cap \mathcal{U}_2$, which is a nonempty open subset of U, we now obtain Claim for every $u \in \mathcal{U}(k)$.

Finally, for such a point u as in Claim, by the containment lemma (Proposition 2.2), $i_u^*(B_U)$ and $i_u^*(V_U)$ have modulus D. Hence, $i_u^*(B_U) \in z^q(X|D, n+1)$ and $i_u^*(V_U) \in z^q_{\mathcal{W}}(X|D, n)$. This finishes the proof. $\square$

Sketch of the proof of Theorem 4.1. Step 1. We first show it when $X = \mathbb{A}^d_k$. Let $K = k(\mathbb{A}^d_k)$ and let $\eta \in X$ be the generic point. To facilitate the proof, as we did previously in §3, using the automorphism $y \mapsto 1/(1-y)$ of $\mathbb{P}^1$ we replace $(\square, \{\infty, 0\})$ by $(\mathbb{A}^1, \{0, 1\})$, and write $\square = \mathbb{A}^1$. We use the homogeneous coordinates $(u_{i,0}; u_{i,1}) \in \overline{\square}^1 = \mathbb{P}^1$, where $y_i = u_{i,1}/u_{i,0}$, then the divisor $F^1_{n,i}$ in the modulus condition is replaced by $F^\infty_{n,i} = \{y_i = \infty\}$ and $F^\infty_n = \sum_{i=1}^n F^\infty_{n,i}$.

For any $g \in \mathbb{A}^d$ and an integer $s > 0$, define $\phi_{g,s} : \mathbb{A}^d_{k(g)}[r] \times_{k(g)} \square^1_{k(g)} \to \mathbb{A}^d_{k(g)}[r]$ by $\phi_{g,s}(\underline{x}, \underline{t}, y) := (\underline{x} + y(t_1^{m_1} \cdots t_r^{m_r})^s g, \underline{t})$, where $k(g)$ is the residue field of g. (N.B. In terms of W. Kai's homotopy, our $g \in \mathbb{A}^d$ corresponds to his $v = (g, 0, \cdots, 0) \in \mathbb{A}^d[r] = \mathbb{A}^{d+r}$.) For any cycle $V \in z^q(X[r]|D_{\underline{m}}, n)$, define $H^*_{g,s}(V) := (\phi_{g,s} \times \mathrm{Id}_{\square^n})^* p^*_{k(g)/k}(V)$, where $p_{k(g)/k} : \mathbb{A}^d_{k(g)}[r] \times \square^n \to \mathbb{A}^d_k[r] \times \square^n$ is the base change.

Using [3, Lemma 1.2], one checks that $H^*_{g,s}(V)$ preserves the face condition for V. Moreover, if $V \in z^q_{\mathcal{W}}(X[r], n)$, then so does $H^*_{g,s}(V)$. When $g = \eta$, another application of [3, Lemma 1.2] shows that $H^*_{g,s}(V)$ intersects with all $W[r] \times F$ properly, where $W \in \mathcal{W}$ and a $F \subset \square^n$ is a face. The argument for proving these face conditions follows the same steps as that of the proof of [17, Lemma 5.5, Case 2] though the present case is slightly different so that we use [3, Lemma 1.2] instead of [3, Lemma 1.1] (see [14, Lemma 3.5] for more detail). On the other hand, we have the following crucial and central assertion due to W. Kai (*cf.* [14, Proposition 3.3]):

Claim: *For each irreducible $V \in z^q(\mathbb{A}^d_k[r]|D_{\underline{m}}, n)$, there is $s(V) \in \mathbb{Z}_{\geq 0}$ such that for any $s > s(V)$ and for any $g \in \mathbb{A}^d$, the cycle $H^*_{g,s}(V)$ has modulus $D_{\underline{m}}$.*
Once it is proven, call the smallest such integer $s(V)$, *the threshold of* V, for simplicity. Here, instead of translations by $g \in \mathbb{A}^d$ used in usual higher Chow groups of $\mathbb{A}^d$ (which correspond to $s = 0$), Kai uses *adjusted* translations as in the definition of $\phi_{g,s}$, so that near the divisors $\{t_i = 0\}$, the effect of adjusted translation is also small, while away from the divisors $\{t_i = 0\}$, the effect of adjusted translation gets larger, so that for a sufficiently large s, this imbues the desired modulus condition into cycles. Note the following elementary fact (*cf.* [14, Lemma 3.2]), which amounts to rewriting the definitions: *Let A be a commutative ring with unity, $\mathfrak{p} \subset A$ a prime ideal, $\zeta \in A$, and $u \in A \setminus \mathfrak{p}$. Then the element ζ/u of $\kappa(\mathfrak{p})$ is integral over $A/\mathfrak{p}$ if and only if there is a homogeneous polynomial $E(a,b) \in A[a,b]$, which is monic in the variable a, with $E(\zeta, b) \in \mathfrak{p}$ in A.*
For each $I \subset \{1, \cdots, n\}$, consider the open subset $U_I \subset \mathbb{A}^d_k \times \overline{\square}^n$ given by the conditions $u_{i,0} \neq 0$ for $i \in I$ and $u_{i,1} \neq 0$ for $i \notin I$. For $i \notin I$, we let $\overline{y}_i = u_{i,0}/u_{i,1} = y_i^{-1}$. Hence, $U_I = \mathrm{Spec}\,(R_I)$, where $R_I := k[\underline{x}, \underline{t}, \{y_i\}_{i \in I}, \{\bar{y}_i\}_{i \notin I}]$, where $\underline{x} = (x_1, \cdots, x_d)$ and $\underline{t} = (t_1, \cdots, t_r)$. On U_I, the divisor F^∞_n used in the definition of the modulus condition is given by the polynomial $\prod_{i \notin I} \overline{y}_i$.
For an irreducible $V \in z^q(\mathbb{A}^d_k[r]|D_{\underline{m}}, n)$, let $\overline{V}$ be its Zariski closure in $\mathbb{A}^d_k[r] \times \overline{\square}^n$. For a given I, the restriction $\overline{V} \cap (\mathbb{A}^d_k[r] \times U_I)$ is given by an ideal of R_I, say, generated by a finite set of polynomials $f^I_\lambda(\underline{x}, \underline{t}, \{y_i\}_{i \in I}, \{\overline{y}_i\}_{i \notin I}) \in R_I$ for $\lambda \in \Lambda_I$.
By the above fact and the assumption that V has the modulus condition, there is a polynomial $E_I(a,b) = E_I(\underline{x}, \underline{t}, \{y_i\}_{i \in I}, \{\overline{y}_i\}_{i \notin I}, a, b) \in R_I[a,b]$, homogeneous in a, b and monic in a, satisfying the condition inside the ring R_I:

$$E_I(\prod_{i \notin I} \overline{y}_i, t^{\underline{m}}) \in \sum_{\lambda \in \Lambda_I} (f^I_\lambda), \text{where } t^{\underline{m}} = t_1^{m_1} \cdots t_r^{m_r}. \tag{4.1}$$

If necessary, by multiplying a power of a to E_I, we may assume $\deg E_I \geq \deg_{\underline{x}} f^I_\lambda$, where deg is the homogeneous degree of E_I in the variables a, b and $\deg_{\underline{x}}$ is the total degree with respect to $\underline{x}$. In doing so, we may further assume that $\deg E_I$ is the same for all subset $I \subset \{1, \cdots, n\}$. For this choice of degrees, we let $s(V) = \deg E_I$. If V is not irreducible, then take the maximum of $s(V_i)$ over all irreducible components V_i of V to define $s(V)$. The heart of the proof is to show that this number satisfies the assertions of Claim, which we do now.
We may assume V is irreducible. For any fixed $s > s(V)$ and $g \in \mathbb{A}^d$, let V' be an irreducible component of $H^*_{g,s}(V)$ and let $\overline{V}'$ be its Zariski closure in $\mathbb{A}^d_\kappa[r] \times \overline{\square}^{n+1}$, where $\kappa = k(g)$. We use the coordinates $(y, y_1, \cdots, y_n) \in \overline{\square}^{n+1}$, and for the first $\overline{\square} = \mathbb{P}^1$, use the homogeneous coordinate $(u_0; u_1)$ so that $y = u_1/u_0$ and $\overline{y} := u_0/u_1 = y^{-1}$. Let $\nu : \overline{V}'^N \to \overline{V}$ be the normalization. Note that whether a divisor is effective or not on $\overline{V}'^N$ is a Zariski local question on $\overline{V}'^N$ (thus on $\overline{V}'$), so we may check the modulus condition Zariski locally

near any point $P \in \overline{V}'$. Fix a point P. Let $I \subset \{1, \cdots, n\}$ be the set points i such that P does not map to $\infty \in \mathbb{P}^1_\kappa$ of the $(i+1)$-th projection $\overline{V}' \hookrightarrow \mathbb{A}^d_\kappa[r] \times \overline{\square}^{n+1} \to \overline{\square}_\kappa = \mathbb{P}^1_\kappa$.
There are two possibilities. In the first case $P \in \mathbb{A}^d_\kappa[r] \times \mathbb{A}^1 \times \overline{\square}^n$, i.e. P does not map to $\infty \in \mathbb{P}^1$ for the first projection to $\overline{\square}_\kappa$, the morphism $p_{\kappa/k} \circ (\phi_{g,s} \times \mathrm{Id}^n_{\square}) : \mathbb{A}^d_\kappa[r] \times \mathbb{A}^1 \times \square^n \to \mathbb{A}^d_k[r] \times \square^n$ extends uniquely to $\mathbb{A}^d_\kappa[r] \times \mathbb{A}^1 \times \overline{\square}^n \to \mathbb{A}^d_k[r] \times \overline{\square}^n$. Thus, by pulling-back the relation (4.1), we obtain in the ring $R_I[y]$,

$$E_I(\underline{x} + y(t^{\underline{m}})^s g, \underline{t}, \{y_i\}_{i \in I}, \{\overline{y}_i\}_{i \notin I}, \prod_{i \notin I} \overline{y}_i, t^{\underline{m}}) \in \\ \in \sum_{\lambda \in \Lambda_I} (f^I_\lambda(\underline{x} + y(t^{\underline{m}})^s g, \underline{t}, \{y_i\}_{i \in I}, \{\overline{y}_i\}_{i \notin I})). \tag{4.2}$$

Here, the polynomials $f^I_\lambda(\underline{x} + y(t^{\underline{m}})^s g, \{y_i\}_{i \in I}, \{\overline{y}_i\}_{i \notin I})$ over $\lambda \in \Lambda_I$ define the underlying closed subscheme of the Zariski closure of $H^*_{g,s}(V)$ restricted on the region $\mathrm{Spec}\,(R_I[y])$. Due to the choice of the degrees of E_I and f^I_λ, the relation (4.2) implies that the rational function $\prod_{i \notin I} \overline{y}_i / t^{\underline{m}}$ is integral using fact. In particular, V' satisfies the modulus condition in a neighborhood of P.
In the remaining case $P \notin \mathbb{A}^d_\kappa[r] \times \mathbb{A}^1 \times \overline{\square}^n$, i.e. P *does* map to $\infty \in \mathbb{P}^1$ for the first projection to $\overline{\square}_\kappa$, we use the affine open chart $\mathrm{Spec}\,(R_I[\overline{y}])$ where $u_1 \neq 0$. The defining ideal of $\overline{V}' \cap \mathrm{Spec}\,(R_I[\overline{y}])$ in the ring $R_I[\overline{y}]$ contains the polynomials $\phi^I_\lambda(\underline{x}, \underline{t}, \overline{y}, \{y_i\}_{i \in I}, \{\overline{y}_i\}_{i \notin I}) := f^I_\lambda(\underline{x} + \frac{1}{\overline{y}}(t^{\underline{m}})^s g, \underline{t}, \{y_i\}_{i \in I}, \{\overline{y}_i\}_{i \notin I}) \cdot \overline{y}^{\deg_{\underline{x}}(f^I_\lambda)}$, where $\lambda \in \Lambda_I$. By expanding the definition of ϕ^I_λ, we see that it is of the form

$$\phi^I_\lambda = \overline{y}^{\deg_{\underline{x}}(f^I_\lambda)} f^I_\lambda(\underline{x}, \underline{t}, \{y_i\}_{i \in I}, \{\overline{y}_i\}_{i \notin I}) + (t^{\underline{m}})^s h, \quad h \in R_I[\overline{y}]. \tag{4.3}$$

Express (4.1) as $E_I(\prod_{i \notin I} \overline{y}_i, t^{\underline{m}}) = \sum_{\lambda \in \Lambda_I} b_\lambda f^I_\lambda$ for some $b_\lambda \in R_I$. Let $c_\lambda := \overline{y}^{s(V) - \deg_{\underline{x}}(f^I_\lambda)} \cdot b_\lambda$ (which is in R_I because $s(V) \geq \deg_{\underline{x}}(f^I_\lambda)$). Then from (4.3),

$$\sum_{\lambda \in \Lambda_I} c_\lambda \phi^I_\lambda = \overline{y}^{s(V)} \cdot E_I(\prod_{i \notin I} \overline{y}_i, t^{\underline{m}}) + (t^{\underline{m}})^s g, \tag{4.4}$$

where (keep in mind that $s \geq s(V)$) the right hand side becomes $(\overline{y} \prod_{i \notin I} \overline{y}_i)^{s(V)} + e_1 \overline{y} (\overline{y} \prod_{i \notin I} \overline{y}_i)^{s(V)-1} t^{\underline{m}} + \cdots + (e_{s(v)} \overline{y}^{s(V)} + (t^{\underline{m}})^{s - s(V)} h) \cdot (t^{\underline{m}})^{s(V)}$, which we write as $E'(\overline{y} \prod_{i \notin I} \overline{y}_i, t^{\underline{m}})$ for a polynomial $E'(a, b) \in R_I[\overline{y}][a, b]$, homogeneous in a, b and monic in a. Thus (4.4) is $\sum_{\lambda \in \Lambda_I} c_\lambda \phi^I_\lambda = E'(\overline{y} \prod_{i \notin I} \overline{y}_i, t^{\underline{m}})$, which implies that the rational function $\overline{y} \prod_{i \notin I} \overline{y}_i / t^{\underline{m}}$ is integral on $\overline{V}' \cap \mathrm{Spec}\,(R_I[\overline{y}])$ using fact. Thus V' also satisfies the modulus condition near P. Combining these two cases, we have now proven Claim.
Now consider the subgroup $z^q_{\mathcal{W}[r],e}(X[r]|D_{\underline{m}}, n)^{\leq s} \subset z^q_{\mathcal{W}[r],e}(X[r]|D_{\underline{m}}, n)$ for $s > 0$, consisting of cycles V with its threshold $s(V) \leq s$ (*cf.* [14, §3.4]). We

deduce

$$\frac{z^q_{\mathcal{W}[r],e}(X[r]|D_{\underline{m}}, n)}{z^q_{\mathcal{W}[r]}(X[r]|D_{\underline{m}}, n)} = \lim_{\to s} \frac{z^q_{\mathcal{W}[r],e}(X[r]|D_{\underline{m}}, n)^{\le s}}{z^q_{\mathcal{W}[r]}(X[r]|D_{\underline{m}}, n)^{\le s}}.$$

Then one has the induced map from $H^*_{\eta,s}$,

$$H^*_{\eta,s} : \frac{z^q_{\mathcal{W}[r],e}(X[r]|D_{\underline{m}}, n)^{\le s}}{z^q_{\mathcal{W}[r]}(X[r]|D_{\underline{m}}, n)^{\le s}} \to \frac{z^q_{\mathcal{W}[r],e}(X_K[r]|D_{\underline{m}}, n+1)}{z^q_{\mathcal{W}[r]}(X_K[r]|D_{\underline{m}}, n+1)},$$

which gives a homotopy between the base change $p^*_{K/k}$ and $H^*_{\eta,s}|_{y=1}$. However, $H^*_{\eta,s}|_{y=1}$ is zero on the quotient, while $p^*_{K/k}$ is injective on homology by Proposition 4.3, after taking $s \to \infty$, so that the map $p^*_{K/k}$ is in fact zero on homology. This means, the quotient $z^q_{\mathcal{W}[r],e}(X[r]|D_{\underline{m}}, n)/z^q_{\mathcal{W}[r]}(X[r]|D_{\underline{m}}, n)$ is acyclic, proving the moving lemma for $X = \mathbb{A}^d_k$.
Step 2. If X is a general smooth affine k-variety of dimension d, we use the standard generic linear projection trick. We choose a closed immersion $X \hookrightarrow \mathbb{A}^N$ for some $N \gg d$ and run the steps of §6 of [17] (with $\mathbb{P}^n$ replaced by $\mathbb{A}^N$ everywhere) *mutatis mutandis* to conclude the proof of the moving lemma for X from that of affine spaces. We leave the details for the reader. □

4.2. Projective method for multivariate additive higher Chow groups. The following theorem generalizes the moving lemma for additive higher Chow groups of smooth projective schemes [17, Theorem 4.1] to a general setting which includes the multivariate additive higher Chow groups.

Theorem 4.4. *Let (S, D) be a smooth quasi-projective k-variety with an effective Cartier divisor. Let X be a smooth projective k-variety. Let $\mathcal{W}$ be a finite collection of locally closed subsets of X. We let $\mathcal{W} \times S := \{W \times S | W \in \mathcal{W}\}$. Then the inclusion $z^q_{\mathcal{W}\times S}(X \times S|X \times D, \bullet) \hookrightarrow z^q(X \times S|X \times D, \bullet)$ is a quasi-isomorphism. In particular, when $\underline{m} = (m_1, \cdots, m_r) \ge 1$, and $(S, D) = (\mathbb{A}^r, D_{\underline{m}})$, the moving lemma holds for multivariate additive higher Chow groups of smooth projective varieties over k.*

Proof. Most arguments of [17, Theorem 4.1] work with minor changes, so we sketch the proof.
Step 1. We first prove the theorem when $X = \mathbb{P}^d_k$. The algebraic group $SL_{d+1,k}$ acts on $\mathbb{P}^d$. Let $K = k(SL_{d+1,k})$. Then there is a K-morphism $\phi : \square^1_K \to SL_{d+1,K}$ such that $\phi(0) = \mathrm{Id}$, and $\phi(\infty) = \eta$, where η is the generic point of $SL_{d+1,k}$. See [17, Lemma 5.4]. For such ϕ, consider the composition H_n of morphisms

$$\mathbb{P}^d \times S \times \square^{n+1}_K \xrightarrow{\mu_\phi} \mathbb{P}^d \times S \times \square^{n+1}_K \xrightarrow{\mathrm{pr}'_K} \mathbb{P}^d \times S \times \square^n_K \xrightarrow{p_{K/k}} \mathbb{P}^d \times S \times \square^n_k,$$

where $\mu_\phi(\underline{x}, s, y_1, \cdots, y_{n+1}) = (\phi(y_1)\underline{x}, s, y_1, \cdots, y_{n+1})$, pr'_K is the projection dropping y_1, and $p_{K/k}$ is the base change. We claim that H^*_n carries $z^q_{\mathcal{W}\times S}(\mathbb{P}^d \times S|\mathbb{P}^d \times D, n)$ to $z^q_{\mathcal{W}\times S}(\mathbb{P}^d_K \times S|\mathbb{P}^d_K \times D, n+1)$, i.e., for an irreducible cycle $Z \in z^q_{\mathcal{W}\times S}(\mathbb{P}^d \times S|\mathbb{P}^d \times S, n)$, we show that $Z' := H^*_n(Z) \in z^q_{\mathcal{W}\times S}(\mathbb{P}^d_K \times S|\mathbb{P}^d_K \times D, n+1)$.

To do so, we first claim that Z' intersects with $W \times S \times F_K$ properly for each $W \in \mathcal{W}$ and each face $F \subset \square^{n+1}$.
(1) In case $F = \{0\} \times F'$ for some face $F' \subset \square^n$, because $\phi(0) = \mathrm{Id}$, we have $Z' \cap (W \times S \times F_K) \simeq Z_K \cap (W \times S \times F'_K)$. Note that $\dim(W \times S \times F_K) = \dim(W \times S \times F'_K)$. Hence, $\mathrm{codim}_{W \times S \times F_K}(Z' \cap (W \times S \times F_K)) = \dim(W \times S \times F_K) - \dim(Z' \cap (W \times S \times F_K)) = \dim(W \times S \times F'_K) - \dim(Z_K \cap (W \times S \times F'_K)) = \dim(W \times S \times F') - \dim(Z \cap (W \times S \times F')) = \mathrm{codim}_{W \times S \times F'}(Z \cap (W \times S \times F')) \geq q$, because $Z \in z^q_{\mathcal{W} \times S}(\mathbb{P}^d \times S|\mathbb{P}^d \times D, n)$.
(2) In case $F = \{\infty\} \times F'$ for some face $F' \subset \square^n$, $\dim(W \times S \times F_K) = \dim(W \times S \times F'_K)$ and $Z' \cap (W \times S \times F_K) \simeq \eta \cdot (Z_K) \cap (W \times S \times F'_K)$, where $SL_{d+1,k}$ acts on $\mathbb{P}^d \times S \times F'$, naturally on $\mathbb{P}^d$ and trivially on $S \times F'$. Let $A := W \times S \times F'$ and $B := Z \cap (\mathbb{P}^d \times S \times F')$. Thus, $\mathrm{codim}_{W \times S \times F_K}(Z' \cap (W \times S \times F_K)) = \dim(W \times S \times F_K) - \dim(Z' \cap (W \times S \times F_K)) = \dim(W \times S \times F'_K) - \dim(\eta \cdot (Z_K) \cap (W \times S \times F'_K)) =^\dagger \dim(A_K) - \dim(\eta \cdot B_K \cap A_K) = \mathrm{codim}_{A_K}(\eta \cdot B_K \cap A_K)$, where $\dagger$ holds because $Z \cap A = B \cap A$. By applying [3, Lemma 1.1] to $G = SL_{d+1,k}$, and the above A, B on $\mathcal{X} := \mathbb{P}^d \times S \times F'$, there is a non-empty open subset $U \subset G$ such that for all $g \in U$, the intersection $(g \cdot A) \cap B$ is proper on $\mathcal{X}$. By shrinking U, we may assume U is invariant under inverse map, so $g = \eta^{-1} \in U$. Thus, $\mathrm{codim}_{A_K}((\eta \cdot B_K) \cap A_K) \geq \mathrm{codim}_{\mathcal{X}_K}(\eta \cdot B_K)$. Since $\mathrm{codim}_{\mathcal{X}_K}(\eta \cdot B_K) = \mathrm{codim}_{\mathcal{X}_K} B_K$ and $\mathrm{codim}_{\mathcal{X}_K} B_K = q$, we get $\mathrm{codim}_{W \times S \times F_K}(Z' \cap (W \times S \times F_K)) = \mathrm{codim}_{A_K}((\eta \cdot B_K) \cap A_K) \geq \mathrm{codim}_{\mathcal{X}_K} B_K = q$.
(3) In case $F = \square \times F'$ for some face $F' \subset \square^n$, the projection $Z' \cap (W \times S \times \square \times F'_K) \to \square_K$ is flat, being a dominant map to a curve, so $\dim(Z' \cap (W \times S \times \square \times F'_K)) = \dim(Z' \cap (W \times S \times \{\infty\} \times F'_K)) + 1$. We also have $\dim(W \times S \times \square \times F'_K) = \dim(W \times S \times \{\infty\} \times F'_K) + 1$. Hence, we deduce $\mathrm{codim}_{W \times S \times F_K}(Z' \cap (W \times S \times F_K)) = \dim(W \times S \times \square \times F_K) - \dim(Z' \cap (W \times S \times \square \times F'_K)) = \mathrm{codim}_{W \times S \times \{\infty\} \times F'_K}(Z' \cap (W \times S \times \{\infty\} \times F'_K)) \geq^\dagger q$, where $\dagger$ follows from case (2). This shows Z' intersects all faces properly.
Now we show that Z' has modulus $\mathbb{P}^d \times D$. We drop all exchange of the factors, for simplicity. For $p : \mathbb{P}^d \to \mathrm{Spec}\,(k)$, we take $V = p(Z)$ on $S \times \square^n$. Because $Z \subset p^{-1}(p(Z)) = \mathbb{P}^r \times V$, we have $Z' = \mu^*_\phi(\square^1_K \times Z) \subset \mu^*_\phi(\mathbb{P}^d \times \square^1_K \times V) = \mathbb{P}^d \times \square^1_K \times V := Z_1$. By Lemma 2.8, V is admissible on $S \times \square^n$. So, $p^*[V] = \mathbb{P}^d \times V$ is admissible on $\mathbb{P}^d \times S \times \square^n$. In particular, $\mathbb{P}^d \times V$ has modulus $\mathbb{P}^d \times D$. Hence, $Z_1 = \mathbb{P}^d \times \square^1_K \times V$ also has modulus $\mathbb{P}^d_K \times D$. Now, $Z' \subset Z_1$ shows that Z' has modulus $\mathbb{P}^d_K \times D$ by Proposition 2.2. Thus, we proved $Z' \in z^q_{\mathcal{W} \times S}(\mathbb{P}^d_K \times S|\mathbb{P}^d_K \times D, n+1)$.
Going back to the proof, one checks that $H^*_\bullet : z^q(\mathbb{P}^d \times S|\mathbb{P}^d \times D, \bullet) \to z^q(\mathbb{P}^d_K \times S|\mathbb{P}^d \times D, \bullet + 1)$ is a chain homotopy satisfying $\partial H^*(Z) + H^*\partial(Z) = Z_K - \eta \cdot (Z_K)$, and the same holds for $z_{\mathcal{W} \times S}$ by a straightforward computation (see [17, Lemma 5.6]). Furthermore, for each admissible Z, we have $\eta \cdot Z_K \in z^q_{\mathcal{W}_K \times S}(\mathbb{P}^d_K \times S|\mathbb{P}^d_K \times D, n)$, by the above proof of proper intersection of Z' with $W \times S \times F_K$, where $F = \{\infty\} \times F'$ for a face $F' \subset \square^n$. Hence, the base change $p^*_{K/k} : z^q(\mathbb{P}^d_k \times S|\mathbb{P}^d_k \times D, \bullet)/z^q_{\mathcal{W} \times S}(\mathbb{P}^d_k \times S|\mathbb{P}^d_k \times D, \bullet) \to z^q(\mathbb{P}^d_K \times S|\mathbb{P}^d_K \times D, \bullet)/z^q_{\mathcal{W}_K \times S}(\mathbb{P}^d_K \times S|\mathbb{P}^d_K \times D, \bullet)$ is homotopic to $\eta \cdot p^*_{K/k}$, which is

zero on the quotient. That is, $p^*_{K/k}$ on the above quotient complex is zero on homology. However, by the spreading argument (Proposition 4.3), $p^*_{K/k}$ is injective on homology. (N.B. We used here an elementary fact that $k(SL_{d+1,k})$ is purely transcendental over k. To check this fact, first note that by definition $k[SL_{d+1,k}] \simeq k[\{T_{i,j} | 1 \leq i, j \leq d+1\}]/(\det(M) - 1)$ for the $(d+1, d+1)$-matrix $M = [T_{ij}]$ consisting of indeterminates $T_{i,j}$ for $1 \leq i, j \leq d+1$. Here by Cramer's rule we can write $\det(M) - 1 = \alpha T_{d+1,d+1} - \beta - 1$, where $\alpha = \det(M_{d+1,d+1})$, $\beta = \sum_{1 \leq j \leq d} (-1)^{d+1+j} \det(M_{d+1,j})$ and M_{ij} is the (i,j)-minor of M. Here both α and β do not have $T_{d+1,d+1}$. Hence $k[SL_{d+1,k}] \simeq k[\{T_{ij} | 1 \leq i, j \leq d+1, (i,j) \neq (d+1, d+1)\}, \frac{\beta+1}{\alpha}]$. Thus, $k(SL_{d+1,k}) \simeq k(\{T_{ij} | 1 \leq i, j \leq d+1, (i,j) \neq (d+1, d+1)\})$, which is purely transcendental over k.) Hence, the quotient complex $z^q(\mathbb{P}^d \times S | \mathbb{P}^d \times D, \bullet)/z^q_{\mathcal{W} \times S}(\mathbb{P}^d \times S | \mathbb{P}^d \times D, \bullet)$ is acyclic, i.e., the moving lemma holds for $(\mathbb{P}^d \times S, \mathbb{P}^d \times D)$, finishing Step 1.

Step 2. Now let X be a general smooth projective variety of dimension d. In this case, we choose a closed immersion $X \hookrightarrow \mathbb{P}^N$ for some $N \gg d$. We now run the linear projection argument of [17, §6] again without any extra argument to deduce the proof of the moving lemma for X from that of the projective spaces. We leave out the details. □

4.3. Contravariant functoriality. The following contravariant functoriality of multivariate additive higher Chow groups is an immediate application of the moving lemma and the proof is identical to that of [17, Theorem 7.1].

Theorem 4.5. *Let $f : X \to Y$ be a morphism of k-varieties, with Y smooth affine or smooth projective. Let $r \geq 1$ and $\underline{m} = (m_1, \cdots, m_r) \geq 1$. Then there exists a pull-back $f^* : \mathrm{CH}^q(Y[r] | D_{\underline{m}}, n) \to \mathrm{CH}^q(X[r] | D_{\underline{m}}, n)$.*
If $g : Y \to Z$ is another morphism with Z smooth affine or smooth projective, then we have $(g \circ f)^* = f^* \circ g^*$.

Remark 4.6. As a special case, when $r = 1$, we have the pull-back map $f^* : \mathrm{TCH}^q(Y, n; m) \to \mathrm{TCH}^q(X, n; m)$.

4.4. The presheaf $\mathcal{TCH}$. For the rest of the section, we concentrate on additive higher Chow groups. Let $m \geq 0$. By Theorem 4.5, we see that $T^q_{n,m} := \mathrm{TCH}^q(-, n; m)$ is a presheaf of abelian groups on the category $\mathbf{SmAff}_k$, but we do not know if it is a presheaf on the categories $\mathbf{Sm}_k$ or $\mathbf{Sch}_k$. However, we can exploit Theorem 4.5 further to define a new presheaf on $\mathbf{Sm}_k$ and $\mathbf{Sch}_k$. The idea of this detour occurred to the authors while working on [20]. We do it for somewhat more general circumstances.

Let $\mathcal{C}$ be a category and $\mathcal{D}$ be a full subcategory. Let F be a presheaf of abelian groups on $\mathcal{D}$, i.e. $F : \mathcal{D}^{\mathrm{op}} \to (\mathrm{Ab})$ is a functor to the category of abelian groups. For each object $X \in \mathcal{C}$, let $(X \downarrow \mathcal{D})$ be the category whose objects are the morphisms $X \to A$ in $\mathcal{C}$, with $A \in \mathcal{D}$, and a morphism from $h_1 : X \to A$ to $h_2 : X \to B$, with $A, B \in \mathcal{D}$, is given by a morphism $g : A \to B$ in $\mathcal{C}$ such that $g \circ h_1 = h_2$. The functor $F : \mathcal{D}^{\mathrm{op}} \to (\mathrm{Ab})$ induces the functor $(X \downarrow \mathcal{D})^{\mathrm{op}} \to (\mathrm{Ab})$ given by $(X \xrightarrow{h} A) \mapsto F(A)$, also denoted by F.

Definition 4.7. Suppose that for each $X \in \mathcal{C}$, the category $(X \downarrow \mathcal{D})$ is cofiltered. Then define $\mathcal{F}(X) := \underset{(X\downarrow\mathcal{D})^{\mathrm{op}}}{\mathrm{colim}} F$.

In particular, when $\mathcal{C} = \mathbf{Sch}_k$ and $\mathcal{D} = \mathbf{SmAff}_k$, one checks that $(X \downarrow \mathbf{SmAff}_k)$ is cofiltered, and for $X \in \mathbf{Sch}_k$, we define $\mathcal{TCH}^q(X, n; m) := \underset{(X\downarrow\mathbf{SmAff}_k)^{\mathrm{op}}}{\mathrm{colim}} T^q_{n,m}$.

Proposition 4.8. *Let $\mathcal{C}$ be a category and $\mathcal{D}$ be a full subcategory such that for each $X \in \mathcal{C}$, the category $(X \downarrow \mathcal{D})$ is cofiltered. Let F be a presheaf of abelian groups on $\mathcal{D}$ and let $\mathcal{F}$ be as in Definition 4.7.*
Let $f : X \to Y$ be a morphism in $\mathcal{C}$. Then for $X \in \mathcal{C}$, the association $X \mapsto \mathcal{F}(X)$ satisfies the following properties:

(1) *There is a canonical homomorphism $\alpha_X : \mathcal{F}(X) \to F(X)$.*
(2) *If $X \in \mathcal{D}$, then α_X is an isomorphism, and $\alpha : \mathcal{F} \to F$ defines an isomorphism of presheaves on $\mathcal{D}$.*
(3) *There is a canonical pull-back $f^* : \mathcal{F}(Y) \to \mathcal{F}(X)$. If $g : Y \to Z$ is another morphism in $\mathcal{C}$, then we have $(g\circ f)^* = f^*\circ g^*$. So, $\mathcal{F}$ is a presheaf of abelian groups on $\mathcal{C}$. In particular, $\mathcal{TCH}^q(-, n; m)$ is a presheaf of abelian groups on $\mathbf{Sch}_k$, which is isomorphic to* $\mathrm{TCH}^q(-, n; m)$ *on* $\mathbf{SmAff}_k$.

Proof. (1) Let $(X \xrightarrow{h} A) \in (X \downarrow \mathcal{D})^{\mathrm{op}}$. By the given assumption, we have the pull-back $h^* : F(A) \to F(X)$. Regarding $F(X)$ as a constant functor on $(X \downarrow \mathcal{D})^{\mathrm{op}}$, this gives a morphism of functors $F \to F(X)$. Taking the colimits over all h, we obtain $\mathcal{F}(X) \to F(X)$, where $\alpha_X = \mathrm{colim}_h h^*$.
(2) When $X \in \mathcal{D}$, the category $(X \downarrow \mathcal{D})^{\mathrm{op}}$ has the terminal object $\mathrm{Id}_X : X \to X$. Hence, the colimit $\mathcal{F}(X)$ is just $F(X)$.
(3) A morphism $f : X \to Y$ in $\mathcal{C}$ defines a functor $f^\sharp : (Y \downarrow \mathcal{D})^{\mathrm{op}} \to (X \downarrow \mathcal{D})^{\mathrm{op}}$ given by $(Y \xrightarrow{h} A) \mapsto (X \xrightarrow{f} Y \xrightarrow{h} A)$. Thus, taking the colimits of the functors induced by F, we obtain $f^* : \mathcal{F}(Y) \to \mathcal{F}(X)$. For another morphism $g : Y \to Z$, that $(g \circ f)^* = f^* \circ g^*$ can be checked easily using the universal property of the colimits.
In the special case when $\mathcal{C} = \mathbf{Sch}_k$ and $\mathcal{D} = \mathbf{SmAff}_k$ with $F = \mathrm{TCH}^q(-, n; m)$, by Theorem 4.5 we know that F is a presheaf on $\mathbf{SmAff}_k$. So, the above general discussion holds. $\square$

Remark 4.9. Since additive higher Chow groups have pull-backs for flat maps (see [16, Lemma 4.7]), it follows that for $X \in \mathbf{Sm}_k$, $\alpha_{(-)}$ defines a map of presheaves $\mathcal{TCH}^q(-, n; m) \to \mathrm{TCH}^q(-, n; m)$ on the small Zariski site X_{Zar} of X. Proposition 4.8(2) says that this map is an isomorphism for affine open subsets of X. Thus, this map of presheaves on X_{Zar} induces an isomorphism of their Zariski sheafifications.

4.5. Moving lemma for smooth semi-local schemes. One remaining objective in Section 4 is to prove the following semi-local variation of Theorem 4.1:

Theorem 4.10. *Let $Y \in \mathbf{SmLoc}_k$. Let $\mathcal{W}$ be a finite set of locally closed subsets of Y. Then the inclusion $\mathrm{Tz}^q_{\mathcal{W}}(Y, \bullet; m) \hookrightarrow \mathrm{Tz}^q(Y, \bullet; m)$ is a quasi-isomorphism.*

We begin with some basic results related to cycles over semi-local schemes. Recall that when A is a ring and $\Sigma = \{p_1, \cdots, p_N\}$ is a finite subset of $\mathrm{Spec}\,(A)$, the localization at Σ is the localization $A \to S^{-1}A$, where $S = \bigcap_{i=1}^N (A \setminus p_i)$. For a quasi-projective k-scheme X and a finite subset Σ of (not necessarily closed) points of X, the localization X_Σ is defined by reducing it to the case when X is affine by the following elementary fact (see [25, Proposition 3.3.36]) that we use often.

Lemma 4.11. *Let X be a quasi-projective k-scheme. Given any finite subset $\Sigma \subset X$ and an open subset $U \subset X$ containing Σ, there exists an affine open subset $V \subset U$ containing Σ.*

For $X \in \mathbf{Sch}_k$ and a point $x \in X$, the open neighborhoods of x form a cofiltered category and we have functorial flat pull-back maps $(j_U^V)^* : \underline{\mathrm{Tz}}^q(V, n; m) \to \underline{\mathrm{Tz}}^q(U, n; m)$ for $j_U^V : U \hookrightarrow V$ in this category.

Lemma 4.12. *Let $X \in \mathbf{Sch}_k$ and let $x \in X$ be a scheme point. Let $Y = \mathrm{Spec}\,(\mathcal{O}_{X,x})$. Then we have $\mathrm{colim}_{x \in U}\, \underline{\mathrm{Tz}}^q(U, n; m) \xrightarrow{\simeq} \underline{\mathrm{Tz}}^q(Y, n; m)$, where the colimit is taken over all open neighborhoods U of x.*

Proof. Replacing x by an affine open neighborhood of $x \in X$, we may assume that X is affine and write $X = \mathrm{Spec}\,(A)$. Let $\mathfrak{p}_x \subset A$ be the prime ideal that corresponds to the point x and let $S := A \setminus \mathfrak{p}_x$, so that $Y = \mathrm{Spec}\,(S^{-1}A)$. To facilitate our proof, using the automorphism $y \mapsto 1/(1-y)$ of $\mathbb{P}^1$, we identify $\square$ with $\mathbb{A}^1$ and take $\{0, 1\} \subset \mathbb{A}^1$ as the faces. So, $X \times B_n = X \times \mathbb{A}^1 \times \mathbb{A}^{n-1} = \mathrm{Spec}\,(A[t, y_1, \cdots, y_{n-1}])$.
Let $\alpha \in \underline{\mathrm{Tz}}^q(Y, n; m)$. We need to find an open subset $U \subset X$ containing x such that the closure of α in $U \times \mathbb{A}^1 \times \mathbb{A}^{n-1}$ is admissible. For this, we may assume α is irreducible, i.e., it is a closed irreducible subscheme $Z \subset Y \times \mathbb{A}^1 \times \mathbb{A}^{n-1}$. Let $\overline{Z}$ be its Zariski closure in $X \times \mathbb{A}^1 \times \mathbb{A}^{n-1}$. Let $\mathfrak{p}$ be the prime ideal of $B := A[t, y_1, \cdots, y_{n-1}]$ such that $V(\mathfrak{p}) = \overline{Z}$.
For the proper intersection with faces, let $\mathfrak{q} \subset B$ be the prime ideal $(y_{i_1} - \epsilon_1, \cdots, y_{i_s} - \epsilon_s)$, where $1 \le i_1 < \cdots < i_s \le n-1$ and $\epsilon_j \in \{0, 1\}$. Let $\mathfrak{P}$ be a minimal prime of $\mathfrak{p} + \mathfrak{q}$. One checks immediately from the behavior of prime ideals under localizations that there is $a \in S$ such that either $\mathfrak{P}B[a^{-1}] = B[a^{-1}]$ or $\mathrm{ht}(\mathfrak{P}B[a^{-1}]) \ge q + s$. This means, over $U_\mathfrak{q} := \mathrm{Spec}\,(A[a^{-1}])$, either the intersection of $\overline{Z}_{U_\mathfrak{q}}$ with $V(\mathfrak{q})$ is empty, or has codimension $\ge q + s$. Applying this argument to all faces, we can take $U_1 := \bigcap_\mathfrak{q} U_\mathfrak{q}$. Then $\overline{Z}_{U_1}$ intersects all faces of $U_1 \times \mathbb{A}^1 \times \mathbb{A}^{n-1}$ properly.
For the modulus condition, let $\nu : \widehat{Z}^N \to \widehat{Z} \hookrightarrow X \times \mathbb{P}^1 \times (\mathbb{P}^1)^{n-1}$ be the normalization composed with the closed immersion of the further Zariski closure $\widehat{Z}$ of $\overline{Z}$. Let $F_n^\infty = \sum_{i=1}^{n-1} \{y_i = \infty\}$ be the divisor at infinity. For an open set $j : U \hookrightarrow X$, the modulus condition of $\overline{Z}_U$ means $(m+1)[j^*\nu^*\{t = 0\}] \le$

$[j^*\nu^*(F_n^\infty)]$ on $\widehat{Z}_U^N$. Note that there exist only finitely many prime Weil divisors $P_1, \cdots, P_\ell$ on $\widehat{Z}^N$ such that $\mathrm{ord}_{P_i}(\nu^*(F_n^\infty) - (m+1)\nu^*\{t=0\}) < 0$. Their images Q_i under the normalization map $\widehat{Z}^N \to \widehat{Z}$ are still irreducible proper closed subsets of $\widehat{Z}$, thus of $X \times \mathbb{P}^1 \times (\mathbb{P}^1)^{n-1}$. Since $Z = \overline{Z}_Y$ has the modulus condition on $Y \times B_n$ by the given assumption, we have $(Y \times \widehat{B}_n) \cap Q_i = \emptyset$ for each $1 \le i \le \ell$. Thus, there is an affine open subset $U_2 \subset X$ containing x such that $(U_2 \times \widehat{B}_n) \cap Q_i = \emptyset$ for each $1 \le i \le \ell$. Now, by construction, $\overline{Z}_{U_2}$ on $U_2 \times B_n$ satisfies the modulus condition. So, taking an affine open subset $U \subset U_1 \cap U_2$ containing x, we have $\overline{Z}_U \in \underline{\mathrm{Tz}}^q(U, n; m)$. That $(\overline{Z}_U)_Y = Z$ is obvious. □

We can extend this colimit description to semi-local schemes:

Lemma 4.13. *Let Y be a semi-local k-scheme obtained by localizing at a finite set Σ of scheme points of a quasi-projective k-variety X. For a cycle Z on $Y \times B_n$, let $\overline{Z}$ be its Zariski closure in $X \times B_n$.*
Then $Z \in \underline{\mathrm{Tz}}^q(Y, n; m)$ if and only if there exists an affine open subset $U \subset X$ containing Σ, such that $\overline{Z}_U \in \underline{\mathrm{Tz}}^q(U, n; m)$, where $\overline{Z}_U$ is the pull-back of $\overline{Z}$ via the open immersion $U \to X$.

Proof. The direction ($\Leftarrow$) is obvious by pulling back via the flat morphism $Y \hookrightarrow U$. For the direction ($\Rightarrow$), by Lemma 4.12, for each $x \in \Sigma$ we have an affine open neighborhood $U_x \subset X$ of x such that $\overline{Z}_{U_x} \in \underline{\mathrm{Tz}}^q(U_x, n; m)$. Take $W = \bigcup_{x \in \Sigma} U_x$. This is an open subset of X containing Σ. By Lemma 2.9, we have $\overline{Z}_W \in \underline{\mathrm{Tz}}^q(W, n; m)$. On the other hand, by Lemma 4.11, there exists an affine open subset $U \subset W$ containing Σ. By taking the flat pull-back via the open immersion $U \hookrightarrow W$, we get $\overline{Z}_U \in \underline{\mathrm{Tz}}^q(U, n; m)$. □

Lemma 4.14. *Let Y be a semi-local integral k-scheme obtained by localizing at a finite set Σ of scheme points of an integral quasi-projective k-scheme X. Let $Z \in \mathrm{Tz}^q(Y, n; m)$, $W \in \mathrm{Tz}^q(Y, n+1; m)$, and let $\overline{Z}$, $\overline{W}$ be their Zariski closures in $X \times B_n$ and $X \times B_{n+1}$, respectively. For every open subset $U \subset X$, the subscript U means the pull-back to U. Then we have the following:*

1. *If $\partial Z = 0$, we can find an affine open subset $U \subset X$ containing Σ such that $\overline{Z}_U \in \mathrm{Tz}^q(U, n; m)$ and $\partial \overline{Z}_U = 0$.*
2. *If $Z = \partial W$, we can find an affine open subset $U \subset X$ containing Σ such that $\overline{Z}_U \in \mathrm{Tz}^q(U, n; m)$, $\overline{W}_U \in \mathrm{Tz}^q(U, n+1; m)$ and $\overline{Z}_U = \partial \overline{W}_U$.*

Proof. Note that (1) is a special case of (2), so we prove (2) only. Let $Z' := \overline{Z} - \partial \overline{W} \in z^q(X \times B_n)$. If Z' is 0 as a cycle, then take $U_0 = X$. If not, let $Z'_1, \cdots, Z'_s$ be the irreducible components of Z'. Since $Z = \partial W$, each component Z'_i has empty intersection with $Y \times B_n$. So, each $\pi((Z'_i)^c)$ is a non-empty open subset of X containing Σ, where $\pi : X \times B_n \to X$ is the projection, which is open. Take $U_0 = \bigcap_{i=1}^s \pi((Z'_i)^c)$.
On the other hand, Lemma 4.13 implies that there exist open sets $U_1, U_2 \subset X$ containing Σ such that $\overline{Z}_{U_1} \in \mathrm{Tz}^q(U_1, n; m)$ and $\overline{W}_{U_2} \in \mathrm{Tz}^q(U_2, n+1; m)$.

Choose an affine open subset $U \subset U_0 \cap U_1 \cap U_2$ containing Σ, using Lemma 4.11. Then part (2) holds over U by construction. $\square$

Proof of Theorem 4.10. We show that the chain map $\mathrm{Tz}^q_{\mathcal{W}}(Y, \bullet; m) \hookrightarrow \mathrm{Tz}^q(Y, \bullet; m)$ is a quasi-isomorphism. Let X be a smooth affine k-variety with a finite subset $\Sigma \subset X$ such that $Y = \mathrm{Spec}\,(\mathcal{O}_{X,\Sigma})$.

For surjectivity on homology, let $Z \in \underline{\mathrm{Tz}}^q(Y, n; m)$ be such that $\partial Z = 0$. Let $\overline{Z}$ be the Zariski closure of Z in $X \times B_n$. Here, $\partial \overline{Z}$ may not be zero, but by Lemma 4.14(1), there exists an affine open subset $U \subset X$ containing Σ such that we have $\partial \overline{Z}_U = 0$, where $\overline{Z}_U$ is the pull-back of $\overline{Z}$ to U. Let $\mathcal{W}_U = \{W_U | W \in \mathcal{W}\}$, where W_U is the Zariski closure of W in U. Then the quasi-isomorphism $\mathrm{Tz}^q_{\mathcal{W}_U}(U, \bullet; m) \hookrightarrow \mathrm{Tz}^q(U, \bullet; m)$ of Theorem 4.1 shows that there are some $C \in \mathrm{Tz}^q(U, n+1; m)$ and $Z'_U \in \mathrm{Tz}^q_{\mathcal{W}_U}(U, n; m)$ such that $\partial C = \overline{Z}_U - Z'_U$. Let $\iota : Y \hookrightarrow U$ be the inclusion. So, by applying the flat pull-back ι^* (which is equivariant with respect to taking faces), we obtain $\partial(\iota^* C) = Z - \iota^* Z'_U$, and here $\iota^* Z'_U \in \mathrm{Tz}^q_{\mathcal{W}}(Y, n; m)$, *i.e.*, Z is equivalent to a member in $\mathrm{Tz}^q_{\mathcal{W}}(Y, n; m)$.

For injectivity on homology, let $Z \in \mathrm{Tz}^q_{\mathcal{W}}(Y, n; m)$ be such that $Z = \partial Z'$ for some $Z' \in \mathrm{Tz}^q(Y, n+1; m)$. Let $\overline{Z}$ and $\overline{Z}'$ be the Zariski closures of Z and Z' on $X \times B_n$ and $X \times B_{n+1}$, respectively. Then by Lemma 4.14(2), there exists a nonempty open affine subset $U \subset X$ containing Σ such that $\overline{Z}_U = \partial \overline{Z}'_U$. Then the quasi-isomorphism $\mathrm{Tz}^q_{\mathcal{W}_U}(U, \bullet; m) \hookrightarrow \mathrm{Tz}^q(U, \bullet; m)$ of Theorem 4.1 shows that there exists $Z'' \in \mathrm{Tz}^q_{\mathcal{W}_U}(U, n+1; m)$ such that $\overline{Z}_U = \partial Z''$. Pulling back via $\iota : Y \hookrightarrow U$ then shows $Z = \partial(\iota^* Z'')$, with $\iota^* Z'' \in \mathrm{Tz}^q_{\mathcal{W}}(Y, n+1; m)$. $\square$

Using an argument identical to Theorem 4.5 (see [17, Theorem 7.1]), we get:

Corollary 4.15. *Let $f : Y_1 \to Y_2$ be a morphism in $\mathbf{Sch}^{\mathrm{ess}}_k$, where $Y_2 \in \mathbf{SmLoc}_k$. Then there is a natural pull-back $f^* : \mathrm{TCH}^q(Y_2, n; m) \to \mathrm{TCH}^q(Y_1, n; m)$.*

5. The Pontryagin product

Let R be a commutative ring and let (A, d_A) be a differential graded algebra over R. Recall that (*left*) differential graded module M over A is a left A-module M with a grading $M = \oplus_{n \in \mathbb{Z}} M_n$ and a differential d_M such that $A_m M_n \subset M_{m+n}$, $d_M(M_n) \subset M_{n+1}$ and $d_M(ax) = d_A(a)x + (-1)^n a d_M(x)$ for $a \in A_n$ and $x \in M$. A homomorphism of differential graded modules $f : M \to N$ over A is an A-module map which is compatible with gradings and differentials.

In this section, we show that the multivariate additive higher Chow groups have a product structure that resembles the Pontryagin product. We construct a differential operator on these groups in the next section and show that the product and the differential operator together turn multivariate additive higher Chow groups groups into a differential graded module over $\mathbb{W}_m \Omega^\bullet_R$ for suitable m, when $X = \mathrm{Spec}\,(R)$ is in $\mathbf{SmAff}^{\mathrm{ess}}_k$. This generalizes the DGA-structure on additive higher Chow groups of smooth projective varieties in [19]. The base field k is perfect in this section.

5.1. Some cycle computations. We generalize some of [19, §3.2.1, 3.2.2, 3.3]. Let (X, D) be a k-scheme with an effective divisor.
Recall that a permutation $\sigma \in \mathfrak{S}_n$ acts naturally on $\square^n$ via $\sigma(y_1, \cdots, y_n) := (y_{\sigma(1)}, \cdots, y_{\sigma(n)})$. This action extends to cycles on $X \times \square^n$ and $X \times \overline{\square}^n$.
Let $n, r \geq 1$ be given. Consider the finite morphism $\chi_{n,r} : X \times \square^n \to X \times \square^n$ given by $(x, y_1, \cdots, y_n) \mapsto (x, y_1^r, y_2, \cdots, y_n)$. Given an irreducible cycle $Z \subset X \times \square^n$, define $Z\{r\} := (\chi_{n,r})_*([Z]) = [k(Z) : k(\chi_{n,r}(Z))] \cdot [\chi_{n,r}(Z)]$. We extend it $\mathbb{Z}$-linearly.

Lemma 5.1. *If Z is an admissible cycle with modulus D, then so is $Z\{r\}$.*

Proof. The proof is almost identical to that of [19, Lemma 3.11], except that the divisor $(m + 1)\{t = 0\}$ there should be replaced by $D \times \overline{\square}^n$. We give its argument for the reader's convenience.
We may assume Z is irreducible. It is enough to show that $\chi_{n,r}(Z)$ is admissible with modulus D. We first check that it satisfies the face condition of Definition 2.3. When $n = 1$, the proper faces of $\square$ are of codimension 1, and for $\epsilon \in \{0, \infty\}$, we have $\partial_1^\epsilon(\chi_{n,r}(Z)) = r\partial_1^\epsilon(Z)$. When $n \geq 2$, for $\epsilon \in \{0, \infty\}$, we have $\partial_1^\epsilon(\chi_{n,r}(Z)) = r\partial_1^\epsilon(Z)$ and $\partial_i^\epsilon(\chi_{n,r}(Z)) = \chi_{n-1,r}(\partial_i^\epsilon(Z))$ if $i \geq 2$. For faces $F \subset \square^n$ of higher codimensions, we consequently have $F\cdot(\chi_{n,r}(Z)) = r(F\cdot Z)$ if F involves the equations $\{y_1 = \epsilon\}$, and $F\cdot(\chi_{n,r}(Z)) = \chi_{n-c,r}(F\cdot Z)$, otherwise, where c is the codimension of F. Since the intersection $F \cdot Z$ is proper, so is $\chi_{n-c,r}(F \cdot Z)$ by induction on the codimension of faces. This shows $\chi_{n,r}(Z)$ satisfies the face condition.
To show that $W := \chi_{n,r}(Z)$ has modulus D, consider the commutative diagram

$$\begin{tikzcd}
\overline{Z}^N \arrow[r, "\nu_Z"] \arrow[d, "\overline{\chi}^N_{n,r}"] & \overline{Z} \arrow[r, hook, "\iota_Z"] \arrow[d, "\overline{\chi}_{n,r}"] & X \times \overline{\square}^n \arrow[d, "\chi_{n,r}"] \\
\overline{W}^N \arrow[r, "\nu_W"] & \overline{W} \arrow[r, hook, "\iota_W"] & X \times \overline{\square}^n,
\end{tikzcd}$$

where $\overline{Z}, \overline{W}$ are the Zariski closures of Z and W in $X \times \overline{\square}^n$ and ν_Z, ν_W are the respective normalizations. The morphisms $\chi_{n,r}$, $\overline{\chi}_{n,r}$ are the natural induced maps, and $\overline{\chi}^N_{n,r}$ is induced by the universal property of normalization. Since Z has modulus D, we have the inequality

$$[\nu_Z^*\iota_Z^*(D \times \overline{\square}^n)] \leq \sum_{i=1}^n [\nu_Z^*\iota_Z^*\{y_i = 1\}]. \tag{5.1}$$

By the definition of $\chi_{n,r}$, we have $\chi_{n,r}^*(D \times \overline{\square}^n) = D \times \overline{\square}^n$, $\chi_{n,r}^*\{y_1 = 1\} \geq \{y_1 = 1\}$, and $\chi_{n,r}^*\{y_i = 1\} = \{y_i = 1\}$ for $i \geq 2$. Hence (5.1) implies that $[\nu_Z^*\iota_Z^*\chi_{n,r}^*(D \times \overline{\square}^n)] \leq \sum_{i=1}^n [\nu_Z^*\iota_Z^*\chi_{n,r}^*\{y_i = 1\}]$. By the commutativity of the diagram, this implies that ${\overline{\chi}^N_{n,r}}^* \left(\sum_{i=1}^n \nu_W^*\iota_W^*\{y_i = 1\} - \nu_W^*\iota_W^*(D \times \overline{\square}^n)\right) \geq 0$.

By Lemma 2.7, this implies $\sum_{i=1}^n \nu_W^* \iota_W^* \{y_i = 1\} - \nu_W^* \iota_W^* (D \times \overline{\square}^n) \geq 0$, which means W has modulus D. This completes the proof. $\square$

Let $n, i \geq 1$. Suppose X is smooth quasi-projective essentially of finite type over k. Let $(x, y_1, \cdots, y_n, y, \lambda)$ be the coordinates of $X \times \overline{\square}^{n+2}$. Consider the closed subschemes V_X^i on $X \times \square^{n+2}$ given by the equation $(1-y)(1-\lambda) = 1-y_1$ if $i = 1$, and $(1-y)(1-\lambda) = (1-y_1)(1+y_1+\cdots+y_1^{i-1} - \lambda(1+y_1+\cdots+y_1^{i-2}))$ if $i \geq 2$.
Let $\widehat{V}_X^i$ be the Zariski closure of V_X^i in $X \times \overline{\square}^{n+2}$. Let $\pi_1 : X \times \overline{\square}^{n+2} \to X \times \overline{\square}^{n+1}$ be the projection that drops y_1, and let $\pi_1' := \pi_1|_{V_X^i}$. As in [19, Lemma 3.12], one sees that π_1' is proper surjective. For an irreducible cycle $Z \subset X \times \square^n$, define (see [19, Definition 3.13]) $\gamma_Z^i := \pi'_{1*}(V_X^i \cdot (Z \times \square^2))$ as an abstract algebraic cycle. One checks that it is also the Zariski closure of $\nu^i(Z \times \square)$, where $\nu^i : X \times \square^n \times \square \to X \times \square^{n+1}$ is the rational map given by $\nu^i(x, y_1, \cdots, y_n, y) = (x, y_2, y_3, \cdots, y_n, y, \frac{y-y_1^i}{y-y_1^{i-1}})$. We extend the definition of γ_Z^i $\mathbb{Z}$-linearly.

Lemma 5.2. *Let $Z \in z^q(X|D, n)$. Then $\gamma_Z^i \in z^q(X|D, n+1)$.*

Proof. Once we have Lemma 5.1, the proof of Lemma 5.2 is very similar to that of [19, Lemma 3.15], except we replace $(m+1)\{t=0\}$ by $D \times \overline{\square}^{n+1}$. We give its argument for the reader's convenience.
We may assume Z is irreducible. To keep track of n, we write $\gamma_{Z,n}^i = \gamma_Z^i$. We first check that it satisfies the face condition of Definition 2.3. Let $\epsilon \in \{0, \infty\}$. Let $F \subset \square^{n+1}$ be a face. If F involves the equation $\{y_j = \epsilon\}$ for $j = n, n+1$, then by direction computations, we see that $\partial_n^0(\gamma_{Z,n}^i) = \sigma \cdot Z, \partial_{n+1}^0(\gamma_{Z,n}^i) = \sigma \cdot (Z\{i\})$ for the cyclic permutation $\sigma = (1, 2, \cdots, n)$, and $\partial_n^\infty(\gamma_{Z,n}^i) = 0$, $\partial_{n+1}^\infty(\gamma_{Z,n}^i) = \sigma \cdot (Z\{i-1\})$. Since Z is admissible with modulus D, so are $Z\{i\}$ and $Z\{i-1\}$ by Lemma 5.1. In particular, all of $\sigma \cdot Z, \sigma \cdot (Z\{i\})$, and $\sigma \cdot (Z\{i-1\})$ intersect all faces properly. Hence $\gamma_{Z,n}^i$ intersects F properly.
In case F does not involve the equations $\{y_j = \epsilon\}$ for $j = n, n+1$, we prove it by induction on $n \geq 1$. By direction calculations, for $j < n$, we have $\partial_j^\epsilon(\gamma_{Z,n}^i) = \gamma^i_{\partial_j^\epsilon Z, n-1}$ so that the dimension of $\partial_j^i(\gamma_{Z,n}^i)$ is at least one less by the induction hypothesis. Repeated applications of this argument for all other defining equations of F then give the result.
It remains to show that γ_Z^i has modulus D. Every irreducible component of γ_Z^i is of the form $W' = \pi_1'(Z')$, where Z' is an irreducible component of $V_X^i \cdot (Z \times \square^2)$. We prove W' has modulus D. Consider the following commutative diagram

$$\begin{tikzcd}
\overline{Z}'^N \arrow[r, "\nu_{Z'}"] \arrow[d, "\pi_1^N"] & \widehat{V}_X^i \arrow[r, hook, "\iota"] \arrow[d, "\overline{\pi}_1'"] \arrow[dr, "\pi_1'"] & X \times \overline{\square}^{n+2} \arrow[d, "\pi_1"] \\
\overline{W}'^N \arrow[r, "\nu"] & \overline{W}' \arrow[r, hook, "\iota_{W'}"] & X \times \overline{\square}^{n+1},
\end{tikzcd}$$

where $\nu_{Z'}$ is the normalization of the Zariski closure $\overline{Z}'$ of Z' in $\widehat{V}^i_X$, ν is the normalization of the Zariski closure $\overline{W}'$ of W' in $X \times \overline{\square}^{n+1}$, and π_1^N, $\overline{\pi}'_1$ are the induced morphisms. We use $(x, y_1, \cdots, y_n, y, \lambda) \in X \times \overline{\square}^{n+2}$ and $(x, y_2, \cdots, y_n, y, \lambda) \in X \times \overline{\square}^{n+1}$ as the coordinates. From the modulus D condition of Z, we deduce

$$\nu^*_{Z'} \iota^*(D \times \overline{\square}^{n+2}) \leq \sum_{j=1}^{n} \nu^*_{Z'} \iota^* \{y_j = 1\}. \tag{5.2}$$

Note that the above does not involve the divisors $\{y = 1\}$ and $\{\lambda = 1\}$. Since V^i_X is an effective divisor on $X \times \square^{n+2}$ defined by the equation $(1 - y_1)(*) = (1 - y)(1 - \lambda)$ for some polynomial $(*)$, we have $[\nu^*_{Z'} \iota^* \{y_1 = 1\}] \leq [\nu^*_{Z'} \iota^* \{y = 1\}] + [\nu^*_{Z'} \iota^* \{\lambda = 1\}]$.
Since the above diagram commutes, from (5.2) we deduce $\pi_1^{N*} \nu^* \iota^*_{W'} (D \times \overline{\square}^{n+1}) \leq \pi_1^{N*} \left(\sum_{j=2}^n \nu^* \iota^*_{W'} \{y_j = 1\} + \{y = 1\} + \{\lambda = 1\} \right)$. Hence by Lemma 2.7, we deduce $\nu^* \iota^*_{W'} (D \times \overline{\square}^{n+1}) \leq \sum_{j=2}^n \nu^* \iota^*_{W'} \{y_j = 1\} + \{y = 1\} + \{\lambda = 1\}$, which means W' has modulus D. This finishes the proof. □

LEMMA 5.3. *Let $n \geq 2$ and let $Z \in z^q(X|D, n)$ such that $\partial_i^\epsilon(Z) = 0$ for all $1 \leq i \leq n$ and $\epsilon \in \{0, \infty\}$. Let $\sigma \in \mathfrak{S}_n$. Then there exists $\gamma_Z^\sigma \in z^q(X|D, n+1)$ such that $Z = (\mathrm{sgn}(\sigma))(\sigma \cdot Z) + \partial(\gamma_Z^\sigma)$.*

Proof. Its proof is almost identical to that of [19, Lemma 3.16], except that we use Lemma 5.2 instead of [19, Lemma 3.15]. We give its argument for the reader's convenience.
First consider the case when σ is the transposition $\tau = (p, p+1)$ for $1 \leq p \leq n-1$. We do it for $p = 1$ only, i.e. $\tau = (1, 2)$. Other cases of τ are similar. Let ξ be the unique permutation such that $\xi \cdot (x, y_1, \cdots, y_{n+1}) = (x, y_n, y_1, y_{n+1}, y_2, \cdots, y_{n-1})$. Consider the cycle $\gamma_Z^\tau := \xi \cdot \gamma_Z^1$, where γ_Z^1 is as in Lemma 5.2. Being a permutation of an admissible cycle, so is this cycle γ_Z^ξ. Furthermore, by direction calculations, we have $\partial_1^\infty(\gamma_Z^\tau) = 0$, $\partial_1^0(\gamma_Z^\tau) = \tau \cdot Z$, $\partial_3^\infty(\gamma_Z^\tau) = 0$ and $\partial_3^0(\gamma_Z^\tau) = Z$. On the other hand, for $\epsilon \in \{0, \infty\}$, $\partial_2^\epsilon(\gamma_Z^\tau)$ is a cycle obtained from $\gamma^1_{\partial_2^\epsilon(Z)}$ by a permutation action. So, it is 0 because $\partial_2^\epsilon(Z) = 0$ by the given assumptions. Similarly for $j \geq 4$, we have $\partial_j^\epsilon(\gamma_Z^\tau) = 0$. Hence $\partial(\gamma_Z^\tau) = Z + \tau \cdot Z$, as desired.
Now let $\sigma \in \mathfrak{S}_n$ be any. By a basic result from group theory, we can express $\sigma = \tau_r \tau_{r-1} \cdots \tau_2 \tau_1$, where each τ_i is a transposition of the form $(p, p+1)$ as considered before. Let $\sigma_0 := \mathrm{Id}$ and $\sigma_\ell := \tau_\ell \tau_{\ell-1} \cdots \tau_1$ for $1 \leq \ell \leq r$. For each such ℓ, by the previous case considered, we have $(-1)^{\ell-1} \sigma_{\ell-1} \cdot Z + (-1)^{\ell-1} \tau_\ell \cdot \sigma_{\ell-1} \cdot Z = \partial((-1)^{\ell-1} \gamma^{\tau_\ell}_{\sigma_{\ell-1} \cdot Z})$. Since $\tau_\ell \cdot \sigma_{\ell-1} = \sigma_\ell$, by taking the sum of the above equations over all $1 \leq \ell \leq r$, after cancellations, we obtain $Z + (-1)^{r-1} \sigma \cdot Z = \partial(\gamma_Z^\sigma)$, where $\gamma_Z^\sigma := \sum_{\ell=1}^r (-1)^{\ell-1} \gamma^{\tau_\ell}_{\sigma_{\ell-1} \cdot Z}$. Since $(-1)^r = \mathrm{sgn}(\sigma)$, we obtain the desired result. □

5.2. Pontryagin product. Let $X \in \mathbf{Sch}_k^{\mathrm{ess}}$ be an equidimensional scheme. For $\underline{m} = (m_1, \cdots, m_r) \geq 1$, let $\mathrm{CH}(X[r]|D_{\underline{m}}) := \oplus_{q,n} \mathrm{CH}^q(X[r]|D_{\underline{m}}, n)$. For $m \geq 1$, we let $\mathrm{TCH}(X; m) = \oplus_{q,n} \mathrm{TCH}^q(X, n; m) = \oplus_{q,n} \mathrm{CH}^q(X[1]|D_{m+1}, n-1)$. The objective of §5.2 is to prove the following result which generalizes [19, §3].

Theorem 5.4. *Let k be a perfect field. Let $m \geq 0$ and let $\underline{m} = (m_1, \cdots, m_r) \geq 1$. Let X, Y be both either in $\mathbf{SmAff}_k^{\mathrm{ess}}$ or in $\mathbf{SmProj}_k$. Then we have the following:*

1. $\mathrm{TCH}(X; m)$ *is a graded commutative algebra with respect to a product* $\wedge_X$.
2. $\mathrm{CH}(X[r]|D_{\underline{m}})$ *is a graded module over* $\mathrm{TCH}(X; |\underline{m}| - 1)$.
3. *For $f : Y \to X$ with $d = \dim Y - \dim X$, $f^* : \mathrm{CH}(X[r]|D_{\underline{m}}) \to \mathrm{CH}(Y[r]|D_{\underline{m}})$ and $f_* : \mathrm{CH}(Y[r]|D_{\underline{m}}) \to \mathrm{CH}(X[r]|D_{\underline{m}})[-d]$ (if f is proper in addition) are morphisms of graded* $\mathrm{TCH}(X; |\underline{m}| - 1)$*-modules.*

The proof requires a series of results and will be over after Lemma 5.13.

Lemma 5.5. *Let $X_1, X_2 \in \mathbf{Sch}_k^{\mathrm{ess}}$. For $i = 1, 2$ and $r_i \geq 1$, let V_i be a cycle on $X_i \times \mathbb{A}^{r_i} \times \square^{n_i}$ with modulus $\underline{m}_i = (m_{i1}, \cdots, m_{ir_i})$, respectively. Then $V_1 \times V_2$, regarded as a cycle on $X_1 \times X_2 \times \mathbb{A}^{r_1+r_2} \times \square^{n_1+n_2}$ after a suitable exchange of factors, has modulus $(\underline{m}_1, \underline{m}_2)$.*

Proof. We may assume that V_1 and V_2 are irreducible. It is enough to show that each irreducible component $W \subset V_1 \times V_2$ has modulus $(\underline{m}_1, \underline{m}_2)$. Let $\iota_i : \overline{V}_i \hookrightarrow X_i \times \mathbb{A}^{r_i} \times \overline{\square}^{n_i}$ be the Zariski closure of V_i, and let $\nu_{\overline{V}_i} : \overline{V}_i^N \to \overline{V}_i$ be the normalization for $i = 1, 2$. Since k is perfect, [16, Lemma 3.1] says that the morphism $\nu := \nu_{\overline{V}_1} \times \nu_{\overline{V}_2} : \overline{V}_1^N \times \overline{V}_2^N \to \overline{V}_1 \times \overline{V}_2 = \overline{V_1 \times V_2}$ is the normalization. Hence, the composite $\overline{W}^N \xrightarrow{\nu_W} \overline{W} \xhookrightarrow{\iota} \overline{V}_1 \times \overline{V}_2$, where $\overline{W}$ is the Zariski closure of W and ν_W is the normalization of $\overline{W}$, factors into $\overline{W}^N \xrightarrow{\iota^N} \overline{V}_1^N \times \overline{V}_2^N \xrightarrow{\nu} \overline{V}_1 \times \overline{V}_2$, where ι^N is the natural inclusion.
Let $(t_1, \cdots, t_{r_1}, t'_1, \cdots, t'_{r_2}, y_1, \cdots, y_{n_1+n_2}) \in \mathbb{A}^{r_1+r_2} \times \overline{\square}^{n_1+n_2}$ be the coordinates. Consider two divisors $D^1 := \sum_{i=1}^{n_1} \{y_i = 1\} - \sum_{j=1}^{r_1} m_{1j} \{t_j = 0\}, D^2 := \sum_{i=n_1+1}^{n_1+n_2} \{y_i = 1\} - \sum_{j=1}^{r_2} m_{2j} \{t'_j = 0\}$. By the modulus conditions satisfied by V_1 and V_2, we have $((\iota_1 \times 1) \circ (\nu_{\overline{V}_1} \times 1))^* D^1 \geq 0$ and $((1 \times \iota_2) \circ (1 \times \nu_{\overline{V}_2}))^* D^2 \geq 0$. Thus, we have $\nu^*(\iota_1 \times \iota_2)^*(D^1 + D^2) \geq 0$ on $\overline{V}_1^N \times \overline{V}_2^N$ so that $(\iota^N)^* \nu^* (\iota_1 \times \iota_2)^*(D^1 + D^2) \geq 0$ on $\overline{W}^N$. Since $\iota \circ \nu_W = \nu \circ \iota^N$, this is equivalent to $\nu_W^* \iota^* (\iota_1 \times \iota_2)^*(D^1 + D^2) \geq 0$, which shows W has modulus $(\underline{m}_1, \underline{m}_2)$. □

Definition 5.6. Let $r \geq 1$ be an integer and define $\mu : X_1 \times \mathbb{A}^1 \times \square^{n_1} \times X_2 \times \mathbb{A}^r \times \square^{n_2} \to X_1 \times X_2 \times \mathbb{A}^r \times \square^{n_1+n_2}$ by $(x_1, t, \{y_j\}) \times (x_2, \{t_i\}, \{y'_j\}) \mapsto (x_1, x_2, \{tt_i\}, \{y_j\}, \{y'_j\})$.

The map μ is flat, but not proper. But, the following generalization of [19, Lemma 3.4] gives a way to take a push-forward:

Proposition 5.7. *Let $V_1 \subset X_1 \times \mathbb{A}^1 \times \square^{n_1}$ and $V_2 \subset X_2 \times \mathbb{A}^r \times \square^{n_2}$ be closed subschemes with moduli m and $\underline{m} \geq 1$, respectively. Then $\mu|_{V_1 \times V_2}$ is finite.*

Proof. Since μ is an affine morphism, the proposition is equivalent to show that $\mu|_{V_1 \times V_2}$ is projective.
Set $X = X_1 \times X_2 \times \square^{n_1+n_2}$. Let $\Gamma \hookrightarrow X_1 \times X_2 \times \mathbb{A}^1 \times \mathbb{A}^r \times \mathbb{A}^r \times \square^{n_1+n_2} = X \times \mathbb{A}^1 \times \mathbb{A}^r \times \mathbb{A}^r$ denote the graph of the morphism μ and let $\overline{\Gamma} \hookrightarrow X \times \mathbb{P}^1 \times (\mathbb{P}^1)^r \times (\mathbb{P}^1)^r = X \times P_1 \times P_2 \times P_3$ be its closure, where $P_1 = \mathbb{P}^1$ and $P_2 = P_3 = (\mathbb{P}^1)^r$. Let p_i be the projection of $X \times \mathbb{P}^1 \times (\mathbb{P}^1)^r \times (\mathbb{P}^1)^r$ to $X \times P_i$ for $1 \leq i \leq 3$. Set $\overline{\Gamma}^0 = p_3^{-1}(X \times \mathbb{A}^r)$. Then $p_3 : \overline{\Gamma}^0 \to X \times \mathbb{A}^r$ is projective.
Using the homogeneous coordinates of $P_1 \times P_2 \times P_3$, one checks easily that $Z := \overline{\Gamma}^0 \setminus \Gamma \subset E \cup (\bigcup_{i=1}^r E_i)$ (the union is taken inside $X \times P_1 \times P_2 \times P_3$), where $E = X \times \{\infty\} \times (\{0\})^r \times \mathbb{A}^r$ and $E_i = X \times \{0\} \times ((\mathbb{P}^1)^{i-1} \times \{\infty\} \times (\mathbb{P}^1)^{r-i}) \times \mathbb{A}^r$.
Let $V = V_1 \times V_2$. Let Γ_V be the graph Γ restricted to V and let $\overline{\Gamma}_V$ be its Zariski closure in $X \times P_1 \times P_2 \times P_3$. Since $p_3 : \overline{\Gamma}^0 \to X \times \mathbb{A}^r$ is projective, so is the map $\overline{\Gamma}_V^0 := \overline{\Gamma}_V \cap \overline{\Gamma}^0 \to X \times \mathbb{A}^r$. So, if we show $\overline{\Gamma}_V^0 \cap Z = \emptyset$, then $V \simeq \Gamma_V = \overline{\Gamma}_V^0$ is projective over $X \times \mathbb{A}^r$, which is the assertion of the proposition.
To show $\overline{\Gamma}_V^0 \cap Z = \emptyset$, consider the projections $X \times P_1 \times P_2 \times P_3 \xrightarrow{p_1} X \times P_1 \xrightarrow{\pi_1} X_1 \times P_1 \times \square^{n_1}$. Since the closure $\overline{V}_1$ has modulus $m \geq 1$ on $X_1 \times P_1 \times \square^{n_1}$, we have $\overline{V}_1 \cap (X_1 \times \{0\} \times \square^{n_1}) = \emptyset$. In particular, $\overline{\Gamma}_V \cap E_i \hookrightarrow (\pi_1 \circ p_1)^{-1}(\overline{V}_1 \cap (X_1 \times \{0\} \times \square^{n_1})) = \emptyset$ for $1 \leq i \leq r$.
To show that $\overline{\Gamma}_V^0 \cap E = \emptyset$, consider the projections $X \times P_1 \times P_2 \times P_3 \xrightarrow{p_2} X \times P_2 \xrightarrow{\pi_2} X_2 \times P_2 \times \square^{n_2}$. Since the closure $\overline{V}_2$ has modulus $\underline{m} \geq 1$ on $X_2 \times P_2 \times \square^{n_2}$, we have $\overline{V}_2 \cap (X_2 \times (\{0\})^r \times \square^{n_2}) = \emptyset$. In particular, $\overline{\Gamma}_V \cap E \hookrightarrow (\pi_2 \circ p_2)^{-1}(\overline{V}_2 \cap (X_2 \times (\{0\})^r \times \square^{n_2})) = \emptyset$. This finishes the proof. $\square$

Lemma 5.8. *Let $X \in \mathbf{Sch}_k^{\text{ess}}$ and let V be a cycle on $X \times \mathbb{A}^1 \times \mathbb{A}^r \times \square^n$ with modulus $(|\underline{m}|, \underline{m})$, where $\underline{m} = (m_1, \cdots, m_r) \geq 1$. Suppose $\mu|_V$ is finite. Then the closed subscheme $\mu(V)$ on $X \times \mathbb{A}^r \times \square^n$ has modulus $\underline{m}$.*

Proof. This is a straightforward generalization of [19, Proposition 3.8] and is a simple application of Lemma 2.7. We skip the detail. We only remark that it is crucial for the proof that the $\mathbb{A}^1$-component of the modulus is at least $|\underline{m}|$. $\square$

Definition 5.9. For any irreducible closed subscheme $V \subset X \times \mathbb{A}^1 \times \mathbb{A}^r \times \square^n$ such that $\mu|_V : V \to \mu(V)$ is finite, where μ is as in Definition 5.6, define $\mu_*(V)$ as the push-forward $\mu_*(V) = \deg(\mu|_V) \cdot [\mu(V)]$. Extend it $\mathbb{Z}$-linearly.
If V_1 is a cycle on $X_1 \times \mathbb{A}^1 \times \square^{n_1}$ and V_2 is a cycle on $X_2 \times \mathbb{A}^r \times \square^{n_2}$ such that $\mu|_{V_1 \times V_2}$ is finite, we define the external product $V_1 \times_\mu V_2 := \mu_*(V_1 \times V_2)$. If $p_i = \dim V_i$, then $\dim(V_1 \times_\mu V_2) = p_1 + p_2$. If $X_1 \times X_2$ is equidimensional and if q_i is the codimension of V_i, then $V_1 \times_\mu V_2$ has codimension $q_1 + q_2 - 1$.

Lemma 5.10. *Let $V_1 \in z^{q_1}(X_1[1]|D_m, n_1)$ and $V_2 \in z^{q_2}(X_2[r]|D_{\underline{m}}, n_2)$ with $X_1, X_2 \in \mathbf{Sch}_k^{\text{ess}}$ and $m, \underline{m} \geq 1$. Then $V_1 \times_\mu V_2$ intersects all faces of $X_1 \times X_2 \times \mathbb{A}^r \times \square^{n_1+n_2}$ properly.*

Proof. We may assume that V_1 and V_2 are irreducible. $V_1 \times V_2$ clearly intersects all faces of $X_1 \times X_2 \times \mathbb{A}^1 \times \mathbb{A}^r \times \square^{n_1+n_2}$ properly. It follows from Proposition 5.7 that $\mu|_{V_1 \times V_2}$ is finite. In this case, the proper intersection property of $\mu(V_1 \times_\mu V_2)$ follows exactly like that of the finite push-forwards of Bloch's higher Chow cycles. $\square$

Corollary 5.11. *Let $X_1, X_2, X_3 \in \mathbf{Sch}_k^{\mathrm{ess}}$ be equidimensional and let $\underline{m} \geq 1$. Then there is a product*

$$\times_\mu : z^{q_1}(X_1[1]|D_{|\underline{m}|}, n_1) \otimes z^{q_2}(X_2[r]|D_{\underline{m}}, n_2) \to \\ \to z^{q_1+q_2-1}((X_1 \times X_2)[r]|D_{\underline{m}}, n_1 + n_2)$$

which satisfies the relation $\partial(\xi \times_\mu \eta) = \partial(\xi) \times_\mu \eta + (-1)^{n_1} \xi \times_\mu \partial(\eta)$. It is associative in the sense that $(\alpha_1 \times_\mu \alpha_2) \times_\mu \beta = \alpha_1 \times_\mu (\alpha_2 \times_\mu \beta)$ for $\alpha_i \in z^{q_i}(X_i[1]|D_{|\underline{m}|}, n_i)$ for $i = 1, 2$ and $\beta \in z^{q_3}(X_3[r]|D_{\underline{m}}, n_3)$. In particular, it induces operations $\times_\mu : \mathrm{CH}^{q_1}(X_1[1]|D_{|\underline{m}|}, n_1) \otimes \mathrm{CH}^{q_2}(X_2[r]|D_{\underline{m}}, n_2) \to \mathrm{CH}^{q_1+q_2-1}((X_1 \times X_2)[r]|D_{\underline{m}}, n_1 + n_2)$.

Proof. The existence of $\times_\mu$ on the level of cycle complexes follows from the combination of Proposition 5.7, Lemma 5.8 and Lemma 5.10. The associativity follows from that of the Cartesian product $\times$ and the product $\mu : \mathbb{A}^1 \times \mathbb{A}^1 \to \mathbb{A}^1$. By definition, one checks $\partial(\xi \times \eta) = \partial(\xi) \times \eta + (-1)^{n_1} \xi \times \partial(\eta)$. So, by applying μ_*, we get the required relation. That $\times_\mu$ descends to the homology follows. $\square$

Definition 5.12. Let $\underline{m} = (m_1, \cdots, m_r) \geq 1$ and let X be in $\mathbf{SmAff}_k^{\mathrm{ess}}$ or in $\mathbf{SmProj}_k$. For cycle classes $\alpha_1 \in \mathrm{CH}^{q_1}(X[1]|D_{|\underline{m}|}, n_1)$ and $\alpha_2 \in \mathrm{CH}^{q_2}(X[r]|D_{\underline{m}}, n_2)$, define the internal product $\alpha_1 \wedge_X \alpha_2$ to be $\Delta_X^*(\alpha_1 \times_\mu \alpha_2)$ via the diagonal pull-back $\Delta_X^* : \mathrm{CH}^{q_1+q_2-1}((X \times X)[r]|D_{\underline{m}}, n_1 + n_2) \to \mathrm{CH}^{q_1+q_2-1}(X[r]|D_{\underline{m}}, n_1 + n_2)$. This map exists by Theorem 4.5 and Corollary 4.15.

Lemma 5.13. *$\wedge_X$ is associative in the sense that $(\alpha_1 \wedge_X \alpha_2) \wedge_X \beta = \alpha_1 \wedge_X (\alpha_2 \wedge_X \beta)$ for $\alpha_1, \alpha_2 \in \mathrm{CH}(X[1]|D_{|\underline{m}|})$ and $\beta \in \mathrm{CH}(X[r]|D_{\underline{m}})$. $\wedge_X$ is also graded-commutative on $\mathrm{CH}(X[1]|D_{|\underline{m}|})$.*

Proof. The associativity holds by Corollary 5.11. For the graded-commutativity, first note by Theorem 3.2 that we can find representatives α_1 and α_2 of the given cycle classes whose codimension 1 faces are all trivial. Let σ be the permutation that sends $(1, \cdots, n_1, n_1 + 1, \cdots, n_1 + n_2)$ to $(n_1 + 1, \cdots, n_1 + n_2, 1, \cdots, n_1)$ so that $\mathrm{sgn}(\sigma) = (-1)^{n_1+n_2}$. It follows from Lemma 5.3 that $\alpha_1 \wedge_X \alpha_2 = (-1)^{n_1+n_2} \alpha_2 \wedge_X \alpha_1 + \partial(W)$ for some admissible cycle W, as desired. $\square$

Proof of Theorem 5.4. The proof of (1) and (2) is just a combination of the above discussion under the observation that $\mathrm{TCH}^q(X, n; m) = \mathrm{CH}^q(X[1]|D_{m+1}, n - 1)$ for $m \geq 0$ and $n \geq 1$. To prove (3) for f^*, consider the commutative diagram

$$\begin{tikzcd} Y[r] \times \square^n \arrow[r, "\Delta_Y"] \arrow[d, "f"] & (Y \times Y)[r] \times \square^n \arrow[d, "f\times f"] & (Y \times Y)[r+1] \times \square^n \arrow[l, "\mu_Y"'] \arrow[d, "f\times f"] \\ X[r] \times \square^n \arrow[r, "\Delta_X"] & (X \times X)[r] \times \square^n & (X \times X)[r+1] \times \square^n. \arrow[l, "\mu_X"'] \end{tikzcd} \tag{5.3}$$

There is a finite set $\mathcal{W}$ of locally closed subsets of X such that $f^* : z^{q_1}_{\mathcal{W}}(X[1]|D_{|\underline{m}|}, \bullet) \to z^{q_1}(Y[1]|D_{|\underline{m}|}, \bullet)$ and $f^* : z^{q_2}_{\mathcal{W}}(X[r]|D_{\underline{m}}, \bullet) \to z^{q_2}(Y[r]|D_{\underline{m}}, \bullet)$ can be defined as taking cycles associated to the inverse images. Moreover, it is enough to consider the product of cycles in $z^{q_1}_{\mathcal{W}}(X[1]|D_{|\underline{m}|}, \bullet)$ and $z^{q_2}_{\mathcal{W}}(X[r]|D_{\underline{m}}, \bullet)$ by the moving lemmas Theorems 4.1 and 4.4. For irreducible cycles $V_1 \in z^{q_1}(X[1]|D_{|\underline{m}|}, n_1)$ and $V_2 \in z^{q_2}(X[r]|D_{\underline{m}}, n_2)$, the map μ_Y is finite when restricted to $f^*(V_1) \times f^*(V_2)$ by Lemma 5.7. In particular, $\mu_Y(f^*(V_1) \times f^*(V_2)) \in z^{q_1+q_2-1}((Y \times Y)[r]|D_{\underline{m}}, n_1 + n_2)$.
Since the right square in the diagram (5.3) is transverse, it follows that $f^*(\mu_X(V_2 \times V_2)) = \mu_Y(f^*(V_1) \times f^*(V_2))$ as cycles. The desired commutativity of the product with f^* now follows from the commutativity of the left square in (5.3) and the composition law of Theorem 4.5.
The proof of (3) for f_* is just the projection formula, whose proof is identical to the one given in [19, Theorem 3.19] in the case when $X_1, X_2 \in \mathbf{SmProj}_k$. $\square$

As applications, we obtain:

Corollary 5.14. *Let X be in $\mathbf{SmAff}^{\rm ess}_k$ or in $\mathbf{SmProj}_k$. Then for $q, n \geq 0$ and $\underline{m} \geq 1$, the group $\mathrm{CH}^q(X[r]|D_{\underline{m}}, n)$ is a $\mathbb{W}_{(|\underline{m}|-1)}(k)$-module.*

Proof. Applying Theorem 5.4 to X and the structure map $X \to \mathrm{Spec}\,(k)$, it follows that $\mathrm{CH}(X[r]|D_{\underline{m}})$ is a graded module over $\mathrm{TCH}(k; |\underline{m}| - 1)$. By Corollary 5.11, this yields a $\mathrm{TCH}^1(k, 1; |\underline{m}| - 1)$-module structure on each $\mathrm{CH}^q(X[r]|D_{\underline{m}}, n)$. The corollary now follows from the fact that there is a ring isomorphism $\mathbb{W}_m(k) \xrightarrow{\sim} \mathrm{TCH}^1(k, 1; m)$ for every $m \geq 1$ by [28, Corollary 3.7]. $\square$

We can explain the homotopy invariance of the groups $\mathrm{CH}^q(X, n)$ in terms of additive higher Chow groups as follows.

Corollary 5.15. *For $X \in \mathbf{Sch}^{\rm ess}_k$ which is equidimensional and for $q, n \geq 0$, we have $\mathrm{CH}^q(X[1]|D_1, n) = 0$.*

Proof. By Corollary 5.11, we have a map $\times_\mu : \mathrm{CH}^1(pt[1]|D_1, 0) \otimes \mathrm{CH}^q(X[1]|D_1, n) \to \mathrm{CH}^q(X[1]|D_1, n)$ and it follows from the definition of $\times_\mu$ that $[1] \times_\mu \alpha = \alpha$ for every $\alpha \in \mathrm{CH}^q(X[1]|D_1, n)$, where $[1] \in \mathrm{CH}^1(pt[1]|D_1, 0)$ is the cycle given by the closed point $1 \in \mathbb{A}^1(k)$. It therefore suffices to show that the homology class of 1 is zero. To do so, we may use the identification $(\square, \{\infty, 0\}) \simeq (\mathbb{A}^1, \{0, 1\})$ given by $y \mapsto 1/(1 - y)$ again. Then the cycle $C \subset \mathbb{A}^2$ given by $\{(t, y) \in \mathbb{A}^2 | ty = 1\}$ is an admissible cycle in $z^1(pt[1]|D_1, 1)$ such that $\partial_1([C]) = [1]$ and $\partial_0([C]) = 0$. $\square$

6. The structure of differential graded modules

In this section, we construct a differential operator on the graded module of §5 of multivariate additive higher Chow groups over the univariate additive higher Chow groups, generalizing [19, §4]. We assume that k is perfect and $\mathrm{char}(k) \neq 2$.

6.1. Differential. Let X be a smooth quasi-projective scheme essentially of finite type over k. Let $r \geq 1$ and let $\underline{m} = (m_1, \cdots, m_r) \geq 1$. Let $(\mathbb{G}_m^r)^\times := \{(t_1, \cdots, t_r) \in \mathbb{G}_m^r \mid t_1 \cdots t_r \neq 1\}$. Consider the morphism $\delta_n : (\mathbb{G}_m^r)^\times \times \square^n \to \mathbb{G}_m^r \times \square^{n+1}$, $(t_1, \cdots, t_r, y_1, \cdots, y_n) \mapsto (t_1, \cdots, t_r, \frac{1}{t_1 \cdots t_r}, y_1, \cdots, y_n)$. It induces $\delta_n : X \times (\mathbb{G}_m^r)^\times \times \square^n \to X \times \mathbb{G}_m^r \times \square^{n+1}$.
Recall a closed subscheme $Z \subset X \times \mathbb{A}^r \times \square^n$ with modulus $\underline{m}$ does not intersect the divisor $\{t_1 \cdots t_r = 0\}$. So, it is closed in $X \times \mathbb{G}_m^r \times \square^n$. For such Z, we define $Z^\times := Z|_{X \times (\mathbb{G}_m^r)^\times \times \square^n}$.

Lemma 6.1. *For a closed subscheme $Z \subset X \times \mathbb{A}^r \times \square^n$ with modulus $\underline{m}$, the image $\delta_n(Z^\times)$ is closed in $X \times \mathbb{G}_m^r \times \square^{n+1}$.*

Proof. It is enough to show that $\delta_n : X \times (\mathbb{G}_m^r)^\times \times \square^n \to X \times \mathbb{G}_m^r \times \square^{n+1}$ is a closed immersion. It reduces to show that the map $(\mathbb{G}_m^r)^\times \to \mathbb{G}_m^r \times (\mathbb{P}^1 \setminus \{1\})$ given by $(t_1, \cdots, t_r) \mapsto (t_1, \cdots, t_r, 1/(t_1 \cdots t_r))$ is a closed immersion. This is obvious because the image coincides with the closed subscheme given by the equation $t_1 \cdots t_r y = 1$, where $(t_1, \cdots, t_r, y) \in \mathbb{G}_m^r \times \square$ are the coordinates. $\square$

Definition 6.2 (*cf.* [19, Definition 4.3]). For a closed subscheme $Z \subset X \times \mathbb{A}^r \times \square^n$ with modulus $\underline{m}$, we write $\delta_n(Z) := \delta_n(Z^\times)$. If Z is a cycle, we define $\delta_n(Z)$ by extending it $\mathbb{Z}$-linearly. We may often write $\delta(Z)$ if no confusion arises.

Lemma 6.3. *Let Z be a cycle on $X \times \mathbb{A}^r \times \square^n$ with modulus $\underline{m}$. Then $\delta_n(Z)$ is a cycle on $X \times \mathbb{A}^r \times \square^{n+1}$ with modulus $\underline{m}$.*

Proof. We may suppose that Z is irreducible. Let $V = \delta_n(Z)$, which is *a priori* closed in $X \times \mathbb{G}_m^r \times \square^{n+1}$. If the closure V' of V in $X \times \mathbb{A}^r \times \square^{n+1}$ has modulus $\underline{m}$, then it does not intersect the divisor $\{t_1 \cdots t_r = 0\}$ of $X \times \mathbb{A}^r \times \square^{n+1}$, so $V = V'$, and V is closed in $X \times \mathbb{A}^r \times \square^{n+1}$ with modulus $\underline{m}$. So, we reduce to show that V' has modulus $\underline{m}$.
Let $\overline{Z}$ and $\overline{V}$ be the Zariski closures of Z and V' in $X \times \mathbb{A}^r \times \overline{\square}^n$ and $X \times \mathbb{A}^r \times \overline{\square}^{n+1}$, respectively. Observe that δ_n extends to $\overline{\delta}_n : X \times \mathbb{A}^r \times \overline{\square}^n \to X \times \mathbb{A}^r \times \overline{\square}^{n+1}$, which is induced from $\mathbb{A}^r \overset{\Gamma}{\to} \mathbb{A}^r \times \overline{\square} \overset{\mathrm{Id} \times \sigma}{\to} \mathbb{A}^r \times \overline{\square}$, where Γ is the graph morphism of the composite $\mathbb{A}^r \to \mathbb{A}^1 \hookrightarrow \overline{\square}$ of the product map followed by the open inclusion, $(t_1, \cdots, t_r) \mapsto (t_1 \cdots t_r) \mapsto (t_1 \cdots t_r; 1)$, while $\sigma : \overline{\square} \to \overline{\square}$ is the antipodal automorphism $(a; b) \mapsto (b; a)$, where $(a; b) \in \overline{\square} = \mathbb{P}^1$ are the homogeneous coordinates. Since Γ is a closed immersion and $\mathrm{Id} \times \sigma$ is an isomorphism, the morphism $\overline{\delta}_n$ is projective. Hence, the dominant map $\delta_n|_{Z^\times} : Z^\times \to V$ induces $\overline{\delta}_n|_{\overline{Z}} : \overline{Z} \to \overline{V}$. In particular, we have a commutative

diagram

$$
\begin{tikzcd}
\overline{Z}^N \arrow[r, "\nu_Z"] \arrow[d, "\widetilde{\delta}_n"] & \overline{Z} \arrow[r, hook, "\iota_Z"] \arrow[d, "\overline{\delta}_n|_{\overline{Z}}"] & X \times \mathbb{A}^r \times \overline{\square}^n \arrow[d, "\overline{\delta}_n"] \\
\overline{V}^N \arrow[r, "\nu_V"] & \overline{V} \arrow[r, hook, "\iota_V"] & X \times \mathbb{A}^r \times \overline{\square}^{n+1},
\end{tikzcd}
\tag{6.1}
$$

where ι_Z, ι_V are the closed immersions, ν_Z, ν_V are normalizations, and $\widetilde{\delta}_n$ is given by the universal property of normalization for dominant maps.

By definition, $\overline{\delta}_n^*\{t_j = 0\} = \{t_j = 0\}$ for $1 \le j \le r$. First consider the case $n \ge 1$. Then $\overline{\delta}_n^* F^1_{n+1,i} = F^1_{n,i-1}$ for $2 \le i \le n+1$. Now, $\widetilde{\delta}_n^* \nu_V^* \iota_V^* (\sum_{i=1}^{n+1} F^1_{n+1,i} - \sum_{j=1}^r m_j\{t_j = 0\}) \ge \widetilde{\delta}_n^* \nu_V^* \iota_V^* (\sum_{i=2}^{n+1} F^1_{n+1,i} - \sum_{j=1}^r m_j \{t_j = 0\}) =^\dagger \nu_Z^* \iota_Z^* \overline{\delta}_n^* (\sum_{i=2}^{n+1} F^1_{n+1,i} - \sum_{j=1}^r m_j\{t_j = 0\}) = \nu_Z^* \iota_Z^* (\sum_{i=2}^{n+1} F^1_{n,i-1} - \sum_{j=1}^r m_j\{t_j = 0\}) = \nu_Z^* \iota_Z^* (\sum_{i=1}^{n} F^1_{n,i} - \sum_{j=1}^r m_j\{t_j = 0\}) \ge^\ddagger 0$, where † holds by the commutativity of (6.1) and ‡ holds as Z has modulus $\underline{m}$. Using Lemma 2.7, we can drop $\widetilde{\delta}_n^*$, i.e., V' has modulus $\underline{m}$.

When $n = 0$, we have for $1 \le j \le r$, $\widetilde{\delta}_0^* \nu_V^* \iota_V^* \{t_j = 0\} = \nu_Z^* \iota_Z^* \overline{\delta}_0^* \{t_j = 0\} = \nu_Z^* \iota_Z^* \{t_j = 0\}$, which is 0 because $\overline{Z} \cap \{t_j = 0\} = \emptyset$. Hence, $\widetilde{\delta}_0^* \nu_V^* \iota_V^* (F^1_{1,1} - \sum_{j=1}^r m_j \{t_j = 0\}) = \widetilde{\delta}_0^* \nu_V^* \iota_V^* F^1_{1,1} \ge 0$. Dropping $\widetilde{\delta}_0^*$, we get V' has modulus $\underline{m}$. □

Proposition 6.4. *Let $Z \in z^q(X[r]|D_{\underline{m}}, n)$. Then $\delta(Z) \in z^{q+1}(X[r]|D_{\underline{m}}, n+1)$. Furthermore, δ and ∂ satisfy the equality $\delta\partial + \partial\delta = 0$.*

Proof. We may assume that Z is an irreducible cycle. Let $\partial^\epsilon_{n,i}$ be the boundary given by the face $F^\epsilon_{n,i}$ on $X \times \mathbb{A}^r \times \square^n$, for $1 \le i \le n$ and $\epsilon = 0, \infty$.

Claim: For $\epsilon = 0, \infty$, (i) $\partial^\epsilon_{n+1,1} \circ \delta_n = 0$, (ii) $\partial^\epsilon_{n+1,i} \circ \delta_n = \delta_{n-1} \circ \partial^\epsilon_{n,i-1}$ for $2 \le i \le n+1$.

For (i), we show that $\delta_n(Z) \cap \{y_1 = \epsilon\} = \emptyset$ for $\epsilon = 0, \infty$. Since $\delta_n(Z) \subset V(t_1 \cdots t_r y_1 = 1)$, we have $\delta_n(Z) \cap \{y_1 = 0\} = \emptyset$. On the other hand, if $\delta_n(Z)$ intersects $\{y_1 = \infty\}$, then some t_i must be zero on Z, i.e., Z intersects $\{t_i = 0\}$ for some $1 \le i \le r$. However, since Z has modulus $\underline{m}$, this can not happen. Thus, $\delta_n(Z) \cap \{y_1 = \infty\} = \emptyset$. This shows (i). For (ii), by the definition of δ_n, the diagram

$$
\begin{tikzcd}
(\mathbb{G}_m^r)^\times \times \square^{n-1} \arrow[r, "\iota^\epsilon_{i-1}"] \arrow[d, "\delta_{n-1}"] & (\mathbb{G}_m^r)^\times \times \square^n \arrow[d, "\delta_n"] \\
\mathbb{G}_m^r \times \square^n \arrow[r, "\iota^\epsilon_i"] & \mathbb{G}_m^r \times \square^{n+1}
\end{tikzcd}
$$

is Cartesian. Thus, $\delta_{n-1}((\iota^*_{i-1}(Z)) = (\iota^\epsilon_i)^*(\delta_n(Z))$ by [6, Proposition 1.7], i.e., (ii) holds. This proves the claim.

By Lemma 6.3, we know $\delta_n(Z)$ has modulus $\underline{m}$. Since Z intersects all faces properly, so does $\delta_n(Z)$ by applying (i) and (ii) of the above claim repeatedly. For $\partial\delta + \delta\partial = 0$, note that $\partial\delta_n(Z) = \sum_{i=1}^{n+1}(-1)^i(\partial^\infty_{n+1,i}\delta_n(Z) - \partial^0_{n+1,i}\delta_n(Z)) =^\dagger$

$\sum_{i=2}^{n+1}(-1)^i(\delta_{n-1}\partial_{n,i-1}^{\infty}(Z) - \delta_{n-1}\partial_{n,i-1}^{0}(Z)) = -\sum_{i=1}^{n}(-1)^i(\delta_{n-1}\partial_{n,i}^{\infty}(Z) - \delta_{n-1}\partial_{n,i-1}^{0}(Z)) = -\delta_{n-1}\sum_{i=1}^{n}(-1)^i(\partial_{n,i}^{\infty}(Z) - \partial_{n,i}^{0}(Z)) = -\delta_{n-1}\circ\partial(Z)$, where † holds by the claim. $\square$

Lemma 6.5 and Corollary 6.6 below, which generalize [19, §4.2], have much simpler proofs than *loc.cit.*

Lemma 6.5. *Let $Z \in z^q(X[r]|D_{\underline{m}}, n)$ be such that $\partial_i^{\epsilon}(Z) = 0$ for $1 \le i \le n$ and $\epsilon = 0, \infty$. Then $2\delta^2(Z)$ is the boundary of an admissible cycle with modulus $\underline{m}$.*

Proof. Note that $\delta^2(Z)$ is an admissible cycle on $X \times \mathbb{A}^r \times \square^{n+2}$ with modulus $\underline{m}$, by Proposition 6.4. For the transposition $\tau = (1, 2)$ on the set $\{1, \cdots, n+2\}$, we have $\tau \cdot \delta^2(Z) = \delta^2(Z)$, by the definition of δ. On the other hand, we have $\tau \cdot \delta^2(Z) = -\delta^2(Z) + \partial(\gamma)$ for some admissible cycle γ, by Lemma 5.3. Hence, we have $-\delta^2(Z) + \partial(\gamma) = \delta^2(Z)$, i.e., $2\delta^2(Z) = \partial(\gamma)$, as desired. $\square$

Corollary 6.6. *Let k be a perfect field of characteristic $\neq 2$ and let X be in $\mathbf{SmAff}_k^{\text{ess}}$ or in $\mathbf{SmProj}_k$. Let $\underline{m} \ge 1$. Then $\delta^2 = 0$ on $\mathrm{CH}^q(X[r]|D_{\underline{m}}, n)$.*

Proof. If $r = \underline{m} = 1$, by Corollary 5.15, there is nothing to prove. So, suppose either $r \ge 2$ or $|\underline{m}| \ge 2$. But, if $r \ge 2$, then we automatically have $|\underline{m}| \ge 2$, so we just consider the latter case.

Given $\alpha \in \mathrm{CH}^q(X[r]|D_{\underline{m}}, n)$, by Theorem 3.2, we can find a representative $Z \in z^q(X[r]|D_{\underline{m}}, n)$ such that $\partial_i^{\epsilon}(Z) = 0$ for $1 \le i \le n$ and $\epsilon = 0, \infty$. Then by Lemma 6.5, we have $2\delta^2(\alpha) = 0$.

On the other hand, by Corollary 5.14, the group $\mathrm{CH}^q(X[r]|D_{\underline{m}}, n)$ is a $\mathbb{W}_{(|\underline{m}|-1)}(k)$-module. As $|\underline{m}| \ge 2$ and $\mathrm{char}(k) \neq 2$, it follows that $2 \in (\mathbb{W}_{(|\underline{m}|-1)}(k))^{\times}$. In particular, $\delta^2(\alpha) = 0$. $\square$

6.2. Leibniz rule. We now discuss the Leibniz rule, generalizing [19, §4.3]. Let $X \in \mathbf{Sch}_k^{\text{ess}}$. Let $(x, t, t_1, \cdots, t_r, y_1, \cdots, y_{n+2}) \in X \times \mathbb{A}^{r+1} \times \square^{n+2}$ be the coordinates. Let $T \subset X \times \mathbb{A}^{r+1} \times \square^{n+2}$ be the closed subscheme defined by the equation $ty_{n+1} = y_{n+2}(tt_1 \cdots t_r y_{n+1} - 1)$.

Definition 6.7 (*cf.* [19, Definition 4.9]). Given a closed subscheme $Z \subset X \times \mathbb{A}^{r+1} \times \square^n$, define $C_Z := T \cdot (Z \times \square^2)$ on $X \times \mathbb{A}^{r+1} \times \square^{n+2}$. This is extended $\mathbb{Z}$-linearly to cycles.

Lemma 6.8. *Let Z be a cycle on $X \times \mathbb{A}^{r+1} \times \square^n$ with modulus $\underline{m} = (m_1, \cdots, m_{r+1})$. Then C_Z has modulus $\underline{m}$ on $X \times \mathbb{A}^{r+1} \times \square^{n+2}$.*

Proof. We may assume Z is irreducible. We show that each irreducible component $V \subset C_Z$ has modulus $\underline{m}$. Let $\overline{Z}$ and $\overline{V}$ be the Zariski closures of Z and V in $X \times \mathbb{A}^{r+1} \times \overline{\square}^n$ and $X \times \mathbb{A}^{r+1} \times \overline{\square}^{n+2}$, respectively. The projection $\mathrm{pr} : X \times \mathbb{A}^{r+1} \times \overline{\square}^{n+2} \to X \times \mathbb{A}^{r+1} \times \overline{\square}^n$ that ignores the last two $\overline{\square}^2$ is projective, while its restriction to $X \times \mathbb{A}^{r+1} \times \square^{n+2}$ maps V into Z. So, pr maps

$\overline{V}$ to $\overline{Z}$, giving a commutative diagram

$$\begin{tikzcd} \overline{V}^N \arrow[r, "\nu_V"] \arrow[d, "\mathrm{pr}^N"] & \overline{V} \arrow[r, hook, "\iota_V"] \arrow[d, "\mathrm{pr}|_{\overline{V}}"] & X \times \mathbb{A}^{r+1} \times \overline{\square}^{n+2} \arrow[d, "\mathrm{pr}"] \\ \overline{Z}^N \arrow[r, "\nu_Z"] & \overline{Z} \arrow[r, hook, "\iota_Z"] & X \times \mathbb{A}^{r+1} \times \overline{\square}^{n}, \end{tikzcd} \tag{6.2}$$

where ι_V and ι_Z are the closed immersions, ν_V and ν_Z are normalizations, and pr^N is induced by the universal property of normalization for dominant maps. The modulus condition for V is now easily verified using the pull-back of the modulus condition for Z on $\overline{Z}^N$ and the fact that $\mathrm{pr}^*\{t_j = 0\} = \{t_j = 0\}$ for all j and $\mathrm{pr}^* F^1_{n,i} = F^1_{n+2,i}$ for all i. □

Corollary 6.9. *Let $X_1, X_2 \in \mathbf{Sch}^{\mathrm{ess}}_k$. Let $V_1 \subset X_1 \times \mathbb{A}^1 \times \square^{n_1}$ and $V_2 \subset X_2 \times \mathbb{A}^r \times \square^{n_2}$ be closed subschemes with moduli $|\underline{m}|$ and $\underline{m}$, respectively with $\underline{m} \geq 1$.*
Under the exchange of factors $X_1 \times \mathbb{A}^1 \times \square^{n_1} \times X_2 \times \mathbb{A}^r \times \square^{n_2} \simeq X_1 \times X_2 \times \mathbb{A}^{r+1} \times \square^n$, where $n = n_1 + n_2$, consider the cycle $C_{V_1 \times V_2}$ on $X_1 \times X_2 \times \mathbb{A}^{r+1} \times \square^{n+2}$. Then $\mu|_{C_{V_1 \times V_2}}$ is finite. In particular, $\mu_*(C_{V_1 \times V_2})$ as in Definition 5.9 is well-defined, and has modulus $\underline{m}$.

Proof. We set $V = V_1 \times V_2$. From the definition of μ, the map $\mu : V \times \square^2 \to X_1 \times X_2 \times \mathbb{A}^r \times \square^{n+2}$ is of the form $\mu|_V \times \mathrm{Id}_{\square^2}$. By Proposition 5.7, the map $\mu|_V$ is finite, thus so is $\mu|_V \times \mathrm{Id}_{\square^2} : V \times \square^2 \to X_1 \times X_2 \times \mathbb{A}^r \times \square^{n+2}$. Hence, its restriction to $C_V = T \cdot (V \times \square^2)$ is also finite. The modulus condition for $\mu_*(C_V)$ follows from Lemmas 5.8 and 6.8. □

Definition 6.10 (*cf.* [19, Definition 4.12]). Let $V_1 \in z^{q_1}(X_1[1]|D_{|\underline{m}|}, n_1)$ and $V_2 \in z^{q_2}(X_2[r]|D_{\underline{m}}, n_2)$ with $X_1, X_2 \in \mathbf{Sch}^{\mathrm{ess}}_k$. Let $n = n_1 + n_2$ and define $V_1 \times_{\mu'} V_2$ be the cycle $\sigma \cdot \mu_*(C_{V_1 \times V_2})$, where $\sigma = (n+2, n+1, \cdots, 1)^2 \in \mathfrak{S}_{n+2}$.

Lemma 6.11. *Let V_1, V_2 be as in Definition 6.10. Then $V_1 \times_{\mu'} V_2 \in z^{q_1+q_2-1}((X_1 \times X_2)[r]|D_{\underline{m}}, n_1 + n_2 + 2)$.*

Proof. By Corollary 6.9, the cycle $\mu_*(C_{V_1 \times V_2})$ has modulus $\underline{m}$, thus so does $W := V_1 \times_{\mu'} V_2$. It remains to prove that W intersects all faces properly. Let $\sigma_{n_1} = (n_1 + 1, n_1, \cdots, 1) \in \mathfrak{S}_{n+1}$. Then by direct calculations, we have

$$\begin{cases} \partial_1^\infty W = \sigma_{n_1}(V_1 \times_\mu \delta(V_2)), \partial_1^0 W = 0, \partial_2^\infty W = \delta(V_1 \times_\mu V_2), \\ \partial_2^0 W = \delta(V_1) \times_\mu V_2, \\ \partial_i^\epsilon W = \begin{cases} \partial_{i-2}^\epsilon(V_1) \times_{\mu'} V_2, & \text{for } 3 \leq i \leq n_1 + 2, \\ V_1 \times_{\mu'} \partial_{i-n_1-2}^\epsilon(V_2), & \text{for } n_1 + 3 \leq i \leq n+2, \end{cases} \quad \epsilon \in \{0, \infty\}. \end{cases} \tag{6.3}$$

Since each V_i is admissible, using (6.3), Lemma 5.10, Proposition 6.4 and induction on the codimension of faces, we deduce that W intersects all faces properly. □

Proposition 6.12. *Let $X_1, X_2 \in \mathbf{Sm}^{\mathrm{ess}}_k$. Let $\xi \in z^{q_1}(X_1[1]|D_{|\underline{m}|}, n_1)$ and $\eta \in z^{q_2}(X_2[r]|D_{\underline{m}}, n_2)$. Let $n = n_1 + n_2$ and $q = q_1 + q_2$. Suppose that* *all codimension one faces of ξ and η vanish. Then in the group $z^{q-1}((X_1 \times X_2)[r]|D_{\underline{m}}, n+1)$, the cycle $\delta(\xi \times_\mu \eta) - \delta\xi \times_\mu \eta - (-1)^{n_1}\xi \times_\mu \delta\eta$ is the boundary of an admissible cycle.*

Proof. By (6.3), for $3 \le i \le n_1 + 2$, we have $\partial_i^\epsilon(\xi \times_{\mu'} \eta) = \partial_{i-2}^\epsilon(\xi) \times_{\mu'} \eta = 0$, while for $n_1+3 \le i \le n+2$, we have $\partial_i^\epsilon(\xi \times_{\mu'} \eta) = \xi \times_{\mu'} \partial_{i-n_1-2}^\epsilon(\eta) = 0$. Hence, $\partial(\xi \times_{\mu'} \eta) = \sum_{i=1}^{n+2}(-1)^i(\partial_i^\infty - \partial_i^0)(\xi \times_{\mu'} \eta) = \delta(\xi \times_\mu \eta) - \{\sigma_{n_1} \cdot (\xi \times_\mu \delta\eta) + \delta\xi \times_\mu \eta\}$ by (6.3) for $i = 1, 2$. Equivalently,

$$\delta(\xi \times_\mu \eta) - \delta\xi \times_\mu \eta - \sigma_{n_1} \cdot (\xi \times_\mu \delta\eta) = \partial(\xi \times_{\mu'} \eta). \tag{6.4}$$

But, for $\xi \times_\mu \delta\eta$, notice that

$$\partial_i^\epsilon(\xi \times_\mu \delta\eta) = \begin{cases} \partial_i^\epsilon \xi \times_\mu \delta\eta = 0, & \text{for } 1 \le i \le n_1, \\ \xi \times_\mu \partial_{i-n_1}^\epsilon(\delta\eta), & \text{for } n_1 + 1 \le i \le n+1, \end{cases} \quad \epsilon \in \{0, \infty\}. \tag{6.5}$$

We have $\partial_1^\epsilon(\delta\eta) = 0$ when $i = n_1 + 1$ by Claim (i) of Proposition 6.4, and $\partial_{i-n_1}^\epsilon(\delta\eta) = \delta(\partial_{i-n_1-1}^\epsilon \eta) = \delta(0) = 0$ when $n_1 + 2 \le i \le n+1$ by Claim (ii) of Proposition 6.4. Hence, $\xi \times_\mu \delta\eta$ is a cycle with trivial codimension 1 faces, so, by Lemma 5.3, for some admissible cycle γ, we have $\sigma_{n_1} \cdot (\xi \times_\mu \delta\eta) = \mathrm{sgn}(\sigma_{n_1})(\xi \times_\mu \delta\eta) + \partial(\gamma) = (-1)^{n_1}\xi \times_\mu \delta\eta + \partial(\gamma)$. Putting this back in (6.4), we obtain $\delta(\xi \times_\mu \eta) - \delta\xi \times_\mu \eta - (-1)^{n_1}\xi \times_\mu \delta\eta = \partial(\xi \times_{\mu'} \eta) - \partial(\gamma)$, as desired. □

The above discussion summarizes as follows:

Theorem 6.13. *Let X be in $\mathbf{SmAff}_k^{\mathrm{ess}}$ or in $\mathbf{SmProj}_k$ over a perfect field k with $\mathrm{char}(k) \neq 2$. Let $r \ge 1$ and $\underline{m} = (m_1, \cdots, m_r) \ge 1$. Then the following hold:*

1. *$(\mathrm{CH}(X[1]|D_{|\underline{m}|}), \wedge_X, \delta)$ forms a commutative differential graded $\mathbb{W}_{(|\underline{m}|-1)}\Omega_k^\bullet$-algebra.*
2. *$(\mathrm{CH}(X[r]|D_{\underline{m}}), \delta)$ forms a differential graded $(\mathrm{CH}(X[1]|D_{|\underline{m}|}), \wedge_X, \delta)$-module.*

In particular, $(\mathrm{CH}(X[r]|D_{\underline{m}}), \delta)$ is a differential graded $\mathbb{W}_{(|\underline{m}|-1)}\Omega_k^\bullet$-module.

Proof. The commutative differential graded algebra structure on $\mathrm{CH}(X[1]|D_{|\underline{m}|})$ and the differential graded module structure on $\mathrm{CH}(X[r]|D_{\underline{m}})$ over $\mathrm{CH}(X[1]|D_{|\underline{m}|})$ follows by combining Theorem 5.4, Corollary 6.6 and Proposition 6.12 using Theorem 3.2.
The structure map $p : X \to \mathrm{Spec}\,(k)$ turns $(\mathrm{CH}(X[1]|D_{|\underline{m}|}), \wedge_X, \delta)$ into a differential graded algebra over $(\mathrm{CH}(pt[1]|D_{|\underline{m}|}), \wedge_{pt}, \delta)$ via p^*. Since $\oplus_{n\ge 0}\mathrm{CH}^{n+1}(pt[1]|D_{|\underline{m}|}, n)$ forms a differential graded sub-algebra of $(\mathrm{CH}(pt[1]|D_{|\underline{m}|}), \wedge_{pt}, \delta)$. The map of commutative differential graded algebras $\mathbb{W}_{(|\underline{m}|-1)}\Omega_k^\bullet \to \oplus_{n\ge 0}\mathrm{CH}^{n+1}(pt[1]|D_{|\underline{m}|}, n)$ (see [28]) finishes the proof of the theorem. □

As a consequence of Theorem 6.13 (use Corollary 5.15 when $|\underline{m}| = 1$), we obtain the following property of multivariate additive higher Chow groups.

Corollary 6.14. *Let $r \geq 1$ and $\underline{m} \geq 1$ and let X be in $\mathbf{SmAff}_k^{\mathrm{ess}}$ or in $\mathbf{SmProj}_k$. Then each $\mathrm{CH}^q(X[r]|D_{\underline{m}}, n)$ is a k-vector space provided* $\mathrm{char}(k) = 0$.

7. Witt-complex structure on additive higher Chow groups

Let k be a perfect field of characteristic $\neq 2$. In this section, a *smooth affine k-scheme* means an object in $\mathbf{SmAff}_k^{\mathrm{ess}}$, i.e., an object of either $\mathbf{SmAff}_k$ or $\mathbf{SmLoc}_k$.

Rülling proved in [28] that the additive higher Chow groups of 0-cycles over $\mathrm{Spec}\,(k)$ form a restricted Witt-complex over k. When X is a smooth projective variety over k, it was proven in [19] that additive higher Chow groups of X form a restricted Witt-complex over k. Our objective is to prove the stronger assertion that the additive higher Chow groups of $\mathrm{Spec}\,(R) \in \mathbf{SmAff}_k^{\mathrm{ess}}$ have the structure of a restricted Witt-complex over R.

Since we exclusively use the case $r = 1$ only, we use the older notations $\mathrm{Tz}^q(X, n; m)$ and $\mathrm{TCH}^q(X, n; m)$ instead of $z^q(X[1]|D_{m+1}, n-1)$ and $\mathrm{CH}^q(X[1]|D_{m+1}, n-1)$. For $X \in \mathbf{Sch}_k^{\mathrm{ess}}$, we let $\mathrm{TCH}(X; m) := \oplus_{n,q}\mathrm{TCH}^q(X, n; m)$ and $\mathrm{TCH}^M(X; m) := \oplus_n \mathrm{TCH}^n(X, n; m)$. The superscript M is for *Milnor*. Let $\mathrm{TCH}(X) := \oplus_m \mathrm{TCH}(X; m)$ and $\mathrm{TCH}^M(X) := \oplus_m \mathrm{TCH}^M(X; m)$. We similarly define $\mathcal{TCH}(X; m)$, $\mathcal{TCH}^M(X; m)$, $\mathcal{TCH}(X)$, and $\mathcal{TCH}^M(X)$ for $X \in \mathbf{Sch}_k$ using Definition 4.7.

7.1. Witt-complex structure over k.

In this section, we show that the additive higher Chow groups for an object of $\mathbf{SmAff}_k^{\mathrm{ess}}$ form a functorial restricted Witt-complex over k. For $r \geq 1$, let $\phi_r : \mathbb{A}^1 \to \mathbb{A}^1$ be the morphism $x \mapsto x^r$, which induces $\phi_r : \mathrm{Spec}\,(R) \times B_n \to \mathrm{Spec}\,(R) \times B_n$. By [19, §5.1, 5.2], we have the Frobenius $F_r : \mathrm{TCH}^q(R, n; rm+r-1) \to \mathrm{TCH}^q(R, n; m)$ and the Verschiebung $V_r : \mathrm{TCH}^q(R, n; m) \to \mathrm{TCH}^q(R, n; rm+r-1)$ given by $F_r = \phi_{r*}$ and $V_r = \phi_r^*$. We also have a natural inclusion $\mathfrak{R} : \mathrm{Tz}^q(R, \bullet; m+1) \to \mathrm{Tz}^q(R, \bullet; m)$ for any $m \geq 1$, which induces $\mathfrak{R} : \mathrm{TCH}^q(R, n; m+1) \to \mathrm{TCH}^q(R, n; m)$, called the *restriction*. Finally, by Theorem 6.13, there is a differential $\delta : \mathrm{Tz}^q(R, \bullet; m) \to \mathrm{Tz}^q(R, \bullet + 1; m)$, which induces $\delta : \mathrm{TCH}^q(R, n; m) \to \mathrm{TCH}^q(R, n+1; m)$.

Theorem 7.1. *Let $X \in \mathbf{SmAff}_k^{\mathrm{ess}}$ and $m \geq 1$. Then $\mathrm{TCH}(X; m)$ is a DGA and $\mathrm{TCH}^M(X; m)$ is its sub-DGA. Furthermore, with respect to the operations $\delta, \mathfrak{R}, F_r, V_r$ in the above together with $\lambda = f^* : \mathbb{W}_m(k) = \mathrm{TCH}^1(k, 1; m) \to \mathrm{TCH}^1(X, 1; m)$ for the structure morphism $f : X \to \mathrm{Spec}\,(k)$, $\mathrm{TCH}(X)$ is a restricted Witt-complex over k and $\mathrm{TCH}^M(X)$ is a restricted sub-Witt-complex over k. These structures are functorial.*

Proof. In [19, Theorem 1.1, Scholium 1.2], it was stated that $\mathrm{TCH}(X; m)$ and $\mathrm{TCH}^M(X; m)$ are DGAs, and that $\mathrm{TCH}(X)$ and $\mathrm{TCH}^M(X)$ are restricted Witt-complexes over k with respect to the above $\delta, \mathfrak{R}, F_r, V_r$, provided the moving lemma holds for X. But this is now shown in Theorems 4.1 and 4.10. We give a very brief sketch of this structure and its functoriality.

The functoriality of the restriction operator $\mathfrak{R}$ recalled above, was stated in [19, Corollary 5.19], which we easily check here: let $f : X \to Y$ be a morphism in $\mathbf{SmAff}_k^{\mathrm{ess}}$ and consider the following commutative diagram:

$$\begin{tikzcd} \mathrm{Tz}^q_{\mathcal{W}}(Y, \bullet; m+1) \arrow[r, "f^*"] \arrow[d, hook] & \mathrm{Tz}^q(X, \bullet; m+1) \arrow[d, hook] \\ \mathrm{Tz}^q_{\mathcal{W}}(Y, \bullet; m) \arrow[r, "f^*"] & \mathrm{Tz}^q(X, \bullet; m), \end{tikzcd}$$

where $\mathcal{W}$ is a finite set of locally closed subsets of Y, and the horizontal maps are chain maps given by the inverse images as in the proof of Theorem 4.5 and Corollary 4.15. The diagram and Theorems 4.1 and 4.10 imply that $f^*\mathfrak{R} = \mathfrak{R}f^*$ because the vertical inclusions induce $\mathfrak{R}$ by definition.
For each $r \geq 1$, the Frobenius F_r and Verschiebung V_r recalled in the above are functorial as proven in [19, Lemmas 5.4, 5.9], and that F_r is a graded ring homomorphism is proven in [19, Corollary 5.6].
Finally, the properties (i), (ii), (iii), (iv), (v) in Section 2.2.2, are all proven in [19, Theorem 5.13], where none requires the projectivity assumption. □

Corollary 7.2. *Let $m \geq 1$ be an integer. Then $\mathcal{TCH}(-; m)$ and $\mathcal{TCH}^M(-; m)$ define presheaves of DGAs on $\mathbf{Sch}_k$, and the pro-systems $\mathcal{TCH}(-)$ and $\mathcal{TCH}^M(-)$ define presheaves of restricted Witt-complexes over k on $\mathbf{Sch}_k$.*

Proof. Let $X \in \mathbf{Sch}_k$. By definition, $\mathcal{TCH}(X; m)$ is the colimit over all $(X \to A) \in (X \downarrow \mathbf{SmAff}_k)^{\mathrm{op}}$ of $\mathrm{TCH}(A; m)$. But the category of DGAs is closed under filtered colimits (see [13]) so that $\mathcal{TCH}(X; m)$ is a DGA. For each morphism $f : X \to Y$ in $\mathbf{Sch}_k$, one checks $f^* : \mathcal{TCH}(Y; m) \to \mathcal{TCH}(X; m)$ is a morphism of DGAs. The other assertions follow easily using Theorem 7.1. □

Before we discuss Witt-complexes over R, we state the following behavior of various operators under finite push-forward maps.

Proposition 7.3. *Let $f : X \to Y$ be a finite map in $\mathbf{SmAff}_k^{\mathrm{ess}}$. Then for $r \geq 1$, we have: (a) $f_*\mathfrak{R} = \mathfrak{R}f_*$; (b) $f_*\delta = \delta f_*$; (c) $f_*F_r = F_r f_*$; (d) $f_*V_r = V_r f_*$.*

Proof. The item (a) is obvious and (b) and (c) follow at once from the fact that these operators are defined as push-forward under closed immersion and finite maps and they preserve the faces. For (d), we consider the commutative diagram

(7.1)
$$\begin{tikzcd} X \times \mathbb{A}^1 \arrow[r, "\mathrm{Id}\times\phi_r"] \arrow[d, "f\times\mathrm{Id}"] & X \times \mathbb{A}^1 \arrow[d, "f\times\mathrm{Id}"] \\ Y \times \mathbb{A}^1 \arrow[r, "\mathrm{Id}\times\phi_r"] & Y \times \mathbb{A}^1. \end{tikzcd}$$

Since this diagram is Cartesian and f as well as ϕ preserve the faces, we conclude from [6, Proposition 1.7] that $f_* \circ \phi_r^* = \phi_r^* \circ f_*$. □

7.2. Witt-complex structure over R. Let $X = \mathrm{Spec}\,(R) \in \mathbf{SmAff}_k^{\mathrm{ess}}$. The objective of this section is to strengthen Theorem 7.1 by showing that $\mathrm{TCH}(X)$ is a restricted Witt-complex over R.

Let $m \geq 1$ be an integer. We first define a group homomorphism $\tau_R : \mathbb{W}_m(R) \to \mathrm{TCH}^1(R, 1; m)$ for any k-algebra R. Recall that the underlying abelian group of $\mathbb{W}_m(R)$ identifies with the multiplicative group $(1+tR[[t]])^\times/(1+t^{m+1}R[[t]])^\times$. For each polynomial $p(t) \in (1 + R[[t]])^\times$, consider the closed subscheme of $\mathrm{Spec}\,(R[t])$ given by the ideal $(p(t))$, and let $\Gamma_{(p(t))}$ be its associated cycle. By definition, $\Gamma_{(p(t))} \cap \{t = 0\} = \emptyset$ so that $\Gamma_{(p(t))} \in \mathrm{Tz}^1(R, 1; m)$. We set $\Gamma_{a,n} = \Gamma_{(1-at^n)}$ for $n \geq 1$ and $a \in R$.

Lemma 7.4. *Let $f(t), g(t)$ be polynomials in $R[t]$, and let $h(t) \in R[t]$ be the unique polynomial such that $(1-tf(t))(1-tg(t)) = 1-th(t)$. Then $\Gamma_{(1-th(t))} = \Gamma_{(1-tf(t))} + \Gamma_{(1-tg(t))}$ in $\mathrm{Tz}^1(R, 1; m)$.*

Proof. This is obvious by $(1 - tf(t))(1 - tg(t)) = 1 - th(t)$. □

Lemma 7.5. *For $n \geq m + 1$, we have $\Gamma_{(1-t^n f(t))} \equiv 0$ in $\mathrm{TCH}^1(R, 1; m)$.*

Proof. Consider the closed subscheme $\Gamma \subset X \times \mathbb{A}^1 \times \square$ given by $y_1 = 1-t^n f(t)$. Let $\nu : \overline{\Gamma}^N \to \overline{\Gamma} \hookrightarrow X \times \mathbb{A}^1 \times \mathbb{P}^1$ be the normalization of the Zariski closure $\overline{\Gamma}$ in $X \times \mathbb{A}^1 \times \mathbb{P}^1$. Since $f(t)t^n = 1 - y_1$ on $\overline{\Gamma}$, we see that $n\nu^*\{t = 0\} \leq \nu^*\{y_1 = 1\}$ on $\overline{\Gamma}^N$. Since $n \geq m + 1$, this shows that Γ satisfies the modulus m condition. Since $\partial_1^\infty(\Gamma) = 0$ and $\partial_1^0(\Gamma) = \Gamma_{(1-t^n f(t))}$ (which is of codimension 1), the cycle Γ is an admissible cycle in $\mathrm{Tz}^1(R, 2; m)$ such that $\partial\Gamma = \Gamma_{(1-t^n f(t))}$. This shows that $\Gamma_{(1-t^n f(t))} \equiv 0$ in $\mathrm{TCH}^1(R, 1; m)$. □

Proposition 7.6. *Let R be a k-algebra. Then the map $\tau_R : (1 + tR[t]) \to \mathrm{Tz}^1(R, 1; m)$ that sends a polynomial $1 - tf(t)$ to $\Gamma_{(1-tf(t))}$, defines a group homomorphism $\tau_R : \mathbb{W}_m(R) \to \mathrm{TCH}^1(R, 1; m)$.*

Proof. Every element $p(t) \in (1 + tR[[t]])^\times$ has a unique expression $p(t) = \prod_{n\geq 1}(1 - a_n t^n)$ for $a_n \in R$. For any such $p(t)$, set $p^{\leq m}(t) = \prod_{n=1}^m (1 - a_n t^n)$. We define $\tau_R(p(t)) = \Gamma_{(p^{\leq m}(t))}$. It follows from Lemmas 7.4 and 7.5 that this map descends to a group homomorphism from $\mathbb{W}_m(R)$. □

Recall from [28, Appendix A] that for each $r \geq 1$, we have the Frobenius $F_r : \mathbb{W}_{rm+r-1}(R) \to \mathbb{W}_m(R)$ and the Verschiebung $V_r : \mathbb{W}_m(R) \to \mathbb{W}_{rm+r-1}(R)$. They are given by $F_r(1 - at^n) = (1 - a^{\frac{r}{s}} t^{\frac{n}{s}})^s$, where $s = \gcd(r, n)$ and $V_r(1 - at^n) = 1 - at^{rn}$. On the other hand, as seen in Section 7.1, we have operations F_r and V_r on $\{\mathrm{TCH}^1(R, 1; m)\}_{m\in\mathbb{N}}$.

Lemma 7.7. *Let R be a k-algebra. Then the maps $\tau_R : \mathbb{W}_m(R) \to \mathrm{TCH}^1(R, 1; m)$ of Proposition 7.6 commute with the F_r and V_r operators on both sides.*

Proof. That $\tau_R V_r = V_r \tau_R$, is easy: we have $V_r(\tau_R(1-at^n)) = V_r(\Gamma_{a,n}) = \Gamma_{a,rn}$, while $\tau_R(V_r(1-at^n)) = \Gamma_{(1-at^{rn})} = \Gamma_{a,rn}$.
That $\tau_R F_r = F_r \tau_R$, is slightly more involved. Recall that $F_r(1-at^n) = (1-a^{\frac{r}{s}}t^{\frac{n}{s}})^s$, where $s = \gcd(r,n)$. Write $n = n's$ and $r = r's$, where $1 = (r', n')$. Hence, we have $\tau_R F_r(1-at^n) = s\Gamma_{a^{\frac{r}{s}}, \frac{n}{s}} = sV_{\frac{n}{s}}(\Gamma_{a^{\frac{r}{s}},1}) = sV_{n'}(\Gamma_{a^{r'},1}) =: \clubsuit$, while $F_r\tau_R(1-at^n) = F_r\Gamma_{a,n} = F_rV_n(\Gamma_{a,1}) =: \heartsuit$.
First observe that when $n = 1$, we have $s = 1$, $r = r'$, $n = n' = 1$, and we have $\heartsuit = F_r(\Gamma_{a,1}) = \Gamma_{a^r,1} = \clubsuit$, so that $\tau_R F_r(1-at) = F_r\tau_R(1-at)$, indeed.
For a general $n \geq 1$, we have $F_rV_n = F_{r'}F_sV_sV_{n'} = F_{r'} \circ (s \cdot \mathrm{Id}) \circ V_{n'} = sF_{r'}V_{n'} =^\dagger sV_{n'}F_{r'}$, where $\dagger$ holds because $(r', n') = 1$. Since $F_{r'}(\Gamma_{a,1}) = \Gamma_{a^{r'},1}$ (by the first case), we have $\heartsuit = F_rV_n(\Gamma_{a,1}) = sV_{n'}F_{r'}(\Gamma_{a,1}) = sV_{n'}(\Gamma_{a^{r'},1}) = \clubsuit$. This shows $\tau_R F_r = F_r\tau_R$. $\square$

Remark 7.8. In the proof of Lemma 7.7, we saw that for $s = (r,n)$,

$$F_r(\Gamma_{a,n}) = s\Gamma_{a^{\frac{r}{s}}, \frac{n}{s}}, \quad V_r(\Gamma_{a,n}) = \Gamma_{a,rn}. \tag{7.2}$$

Proposition 7.9. *For $X = \mathrm{Spec}\,(R) \in \mathbf{SmAff}_k^{\mathrm{ess}}$, the maps $\tau_R : \mathbb{W}_m(R) \to \mathrm{TCH}^1(R,1;m)$ form a morphism of pro-rings that commutes with F_r and V_r for $r \geq 1$.*

Proof. It is clear from the definition of τ_R in Proposition 7.6 that it commutes with $\mathfrak{R}$. We saw that τ_R commutes with F_r and V_r in Lemma 7.7. So, we only need to show that τ_R respects the products. By [2, Proposition (1.1)], it is enough to prove that for $a, b \in R$ and $u, v \geq 1$,

$$\Gamma_{a,u} \wedge \Gamma_{b,v} = w\Gamma_{a^{\frac{v}{w}}b^{\frac{u}{w}}, \frac{uv}{w}} \quad \text{in } \mathrm{TCH}^1(R,1;m), \tag{7.3}$$

where $w = \gcd(u,v)$ and $\wedge = \wedge_X$ is the product structure on the ring $\mathrm{TCH}^1(R,1;m)$ as in Theorem 7.1.
Step 1. First, consider the case when $u = v = 1$, *i.e.*, we prove $\Gamma_{a,1} \wedge \Gamma_{b,1} = \Gamma_{ab,1}$. Recall that $\wedge$ is defined as the composition $\Delta^* \circ \mu_* \circ \times$ in

$$X \times \mathbb{A}^1 \times X \times \mathbb{A}^1 \xrightarrow{\mu} X \times X \times \mathbb{A}^1 \xleftarrow{\Delta} X \times \mathbb{A}^1.$$

Under the identification $X \times X \simeq \mathrm{Spec}\,(R \otimes_k R)$, we have $\mu_*(\Gamma_{a,1} \times \Gamma_{b,1}) = \Gamma_{(a\otimes 1)(1\otimes b),1}$, and $\Delta^*(\Gamma_{(a\otimes 1)(1\otimes b),1}) = \Gamma_{ab,1}$, because Δ is given by the multiplication $R \otimes_k R \to R$. This proves (7.3) for Step 1.
For the following remaining two steps, we use the projection formula: $x \wedge V_s(y) = V_s(F_s(x) \wedge y)$, which we can use by Theorem 7.1.
Step 2. Consider the case when $v = 1$, but $u \geq 1$ is any integer. We apply the projection formula to $x = \Gamma_{b,1}$ and $y = \Gamma_{a,1}$ with $s = u$. Since $\mathrm{TCH}^1(R,1;m)$ is a commutative ring, by the projection formula, we get $V_u(\Gamma_{a,1}) \wedge \Gamma_{b,1} = V_u(\Gamma_{a,1} \wedge F_u(\Gamma_{b,1}))$. Here, the left hand side is $\Gamma_{a,u} \wedge \Gamma_{b,1}$ by eqn:FV identity, while the right hand side is $=^1 V_u(\Gamma_{a,1} \wedge \Gamma_{b^u,1}) =^2 V_u(\Gamma_{ab^u,1}) =^3 \Gamma_{ab^u,u}$, where $=^1$ and $=^3$ hold by (7.2) and $=^2$ holds by Step 1. This proves (7.3) for Step 2.
Step 3. Finally, let $u, v \geq 1$ be any integers. Let $w = \gcd(u,v)$. We again apply the projection formula to $x = V_u(\Gamma_{a,1})$, $y = \Gamma_{b,1}$, $s = v$, so that $V_u(\Gamma_{a,1}) \wedge$

$V_v(\Gamma_{b,1}) = V_v(F_v(V_u(\Gamma_{a,1}))\wedge\Gamma_{b,1})$. Its left hand side coincides with that of (7.3) by (7.2). Its right hand side is $=^1 V_v(F_v(\Gamma_{a,u}) \wedge \Gamma_{b,1}) =^2 V_v(w\Gamma_{a\frac{v}{w},\frac{u}{w}} \wedge \Gamma_{b,1})$, where $=^1$ and $=^2$ hold by (7.2). But, Step 2 says that $\Gamma_{a\frac{v}{w},\frac{u}{w}} \wedge \Gamma_{b,1} = \Gamma_{a\frac{v}{w}b\frac{u}{w},\frac{u}{w}}$ so that $V_v(w\Gamma_{a\frac{v}{w},\frac{u}{w}} \wedge \Gamma_{b,1}) = wV_v(\Gamma_{a\frac{v}{w}b\frac{u}{w},\frac{u}{w}}) =^\dagger w\Gamma_{a\frac{v}{w}b\frac{u}{w},\frac{uv}{w}}$, where $= \dagger$ holds by (7.2). This last expression is the right hand side of (7.3). Thus, we obtain the equality (7.3) and this finishes the proof. □

Theorem 7.10. *For* $\mathrm{Spec}\,(R) \in \mathbf{SmAff}_k^{\mathrm{ess}}$*,* $\mathrm{TCH}(R)$ *is a restricted Witt-complex over* R*, and its sub-pro-system* $\mathrm{TCH}^M(R)$ *is a restricted sub-Witt-complex over* R*.*

Proof. As saw in the proof of Theorem 7.1, we already have the restriction $\mathfrak{R}$, the differential δ, the Frobenius F_r and the Verschiebung V_r defined by the same formulas. Furthermore, by Proposition 7.9, now we have ring homomorphisms $\lambda = \tau_R : \mathbb{W}_m(R) \to \mathrm{TCH}^1(R, 1; m)$ for $m \geq 1$. The properties (i), (ii), (iii), (iv) in Section 2.2.2 are independent of the choice of the ring, so that what we checked in Theorem 7.1 still work. To prove the theorem, the only thing left to be checked is the property (v) that for all $a \in R$ and $r \geq 1$,

$$F_r\delta\tau_R([a]) = \tau_R([a]^{r-1})\delta\tau_R([a]), \tag{7.4}$$

where we have shrunk the product notation $\wedge$ and taken the ring homomorphism λ to be τ_R. To check this, we identify $\mathbb{W}_m(R)$ with $(1 + tR[[t]])^\times/(1 + t^{m+1}R[[t]])^\times$.

If $a = 0$, then $\tau_R([a]) = \Gamma_{(1-0\cdot t)} = \emptyset$. So, both sides of (7.4) are zero.

If $a = 1$, then $\tau_R([a]) = \tau_R(1-t) = \Gamma_{(1-t)}$. But, in our definition of δ, to compute it, we should first restrict the cycle $\Gamma_{(1-t)} \subset \mathrm{Spec}\,(R) \times \mathbb{G}_m$ onto $\mathrm{Spec}\,(R) \times (\mathbb{G}_m \setminus \{1\})$, which becomes empty. Hence, $\delta\tau_R([a]) = \delta\Gamma_{(1-t)} = 0$, so again both sides of (7.4) are zero.

Let $a \in R \setminus \{0, 1\}$. Then $\tau_R([a]) = \Gamma_{(1-at)} \subset \mathrm{Spec}\,(R) \times \mathbb{A}^1$, and $\delta\tau_R([a])$ is given by the ideal $(1 - at, 1 - ty_1)$ in $R[t, y_1]$. Since t is not a zero-divisor in $R[t, y_1]$, we have $(1 - at, 1 - ty_1) = (1 - at, y_1 - a)$ as ideals. Hence, $F_r\delta\tau_R([a])$ is given by the ideal $(1 - a^r t, y_1 - a)$ in $R[t, y_1]$. On the other hand,

$$\begin{aligned} \tau_R([a]^{r-1})\delta\tau_R([a]) &= \Gamma_{(1-a^{r-1}t)} \wedge \mathrm{Spec}\left(\tfrac{R[t,y_1]}{(1-at,y_1-a)}\right) \\ = \quad \Delta^*\left(\tfrac{(R\otimes_k R)[t,y_1]}{(1-(a^{r-1}\otimes 1)(1\otimes a),y_1-(1\otimes a))}\right) &=^\dagger \mathrm{Spec}\left(\tfrac{R[t,y_1]}{(1-a^r t,y_1-a)}\right), \end{aligned} \tag{7.5}$$

where † holds because Δ is induced by the product homomorphism $R\otimes_k R \to R$. Hence, both hand sides of (7.4) coincide. This completes the proof. □

Theorem 7.11. *For* $\mathrm{Spec}\,(R) \in \mathbf{SmAff}_k^{\mathrm{ess}}$ *and* $n, m \geq 1$*, there is a unique homomorphism* $\tau_{n,m}^R : \mathbb{W}_m\Omega_R^{n-1} \to \mathrm{TCH}^n(R, n; m)$ *that defines a morphism of restricted Witt-complexes over* R*,* $\{\tau_{\bullet,m}^R : \mathbb{W}_m\Omega_R^{\bullet-1} \to \mathrm{TCH}^\bullet(R, \bullet; m)\}_m$*, such that* $\tau_{1,m}^R = \tau_R$*.*

Proof. The theorem follows from Theorem 7.10 and [28, Proposition 1.15]. We have $\tau_{1,m}^R = \tau_R$ because the map λ of §2.2.2 is given by τ_R in Theorem 7.10. □

We have shown in Propositions 7.6 and 7.9 that τ_R is a group homomorphism for any k-algebra R and is a ring homomorphism if R is smooth. Here, we provide the following information on τ_R.

Theorem 7.12. *Let R be an integral domain which is an essentially of finite type k-algebra. Then τ_R is injective. It is an isomorphism if R is a UFD.*

Proof. Let $K := \mathrm{Frac}(R)$ and $\iota : R \hookrightarrow K$ be the inclusion. This induces a commutative diagram

$$\begin{tikzcd} \mathbb{W}_m(R) \arrow[r, "\mathbb{W}_m(\iota)"] \arrow[d, "\tau_R"'] & \mathbb{W}_m(K) \arrow[d, "\tau_K"', "\simeq"] \\ \mathrm{TCH}^1(R,1;m) \arrow[r] & \mathrm{TCH}^1(K,1;m), \end{tikzcd}$$

where the bottom map is the flat pull-back via $\mathrm{Spec}\,(K) \to \mathrm{Spec}\,(R)$, and τ_K is the isomorphism by [28, Corollary 3.7]. Since $\mathbb{W}_m(\iota)$ is clearly injective (see [28, Properties A.1.(i)]), it follows that τ_R is injective, too.
Suppose now R is a UFD and V is an irreducible admissible cycle in $\mathrm{Tz}^1(R,1;m)$. Then we must have $(I(V),t) = R[t]$, where $I(V)$ is the ideal of V. Since $R[t]$ is a UFD, using basic commutative algebra, one checks that $I(V) = (1 - tf(t))$ for some non-zero polynomial $f(t) \in R[t]$. In particular, the map τ_R is surjective and hence an isomorphism. □

7.3. Étale descent. Finally:

Proof of Theorem 1.4. By Corollary 5.15, we can assume $|\underline{m}| \geq 2$. We set $Y = X/G$, $\lambda = |G|$ and consider the diagram

$$\begin{tikzcd} G \times X \arrow[r, "\gamma"] \arrow[d, "p"'] & X \arrow[d, "f"] \\ X \arrow[r, "f"'] & Y, \end{tikzcd} \tag{7.6}$$

where γ is the action map and p is the projection. Since G acts freely on X, this square is Cartesian and f is étale of degree λ. By [6, Proposition 1.7], we have $f^* \circ f_* = p_* \circ \gamma^* : \mathrm{CH}^q(X[r]|D_{\underline{m}}, n) \to \mathrm{CH}^q(X[r]|D_{\underline{m}}, n)$.
Since f is G-equivariant with respect to the trivial G-action on Y, we see that f^* induces a map $f^* : \mathrm{CH}^q(Y[r]|D_{\underline{m}}, n) \to \mathrm{CH}^q(X[r]|D_{\underline{m}}, n)^G$. Moreover, it follows from [21, Theorem 3.12] that $f_* \circ f^*$ is multiplication by λ.
On the other hand, it follows easily from the action map γ that $p_* \circ \gamma^*(\alpha) = \sum_{g \in G} g^*(\alpha)$. In particular, $p_* \circ \gamma^*(\alpha) = \lambda \cdot \alpha$ if $\alpha \in \mathrm{CH}^q(X[r]|D_{\underline{m}}, n)^G$.
Since $\lambda \in k^\times$ and the Teichmüller map is multiplicative with $|\underline{m}| \geq 2$, we see that $\lambda \in (\mathbb{W}_{(|\underline{m}|-1)}(k))^\times$. We conclude from Theorem 5.4(3) and Corollary 5.14 that the composite $\mathrm{CH}^q(Y[r]|D_{\underline{m}}, n) \xrightarrow{f^*} \mathrm{CH}^q(X[r]|D_{\underline{m}}, n)^G \xrightarrow{\lambda^{-1} f_*} \mathrm{CH}^q(Y[r]|D_{\underline{m}}, n)$ yields the desired isomorphism. □

Acknowledgments. The authors thank Wataru Kai for letting them know about his moving lemma for affine schemes. The authors also feel very grateful to the referee whose careful and detailed comments had improved the paper and the editor of the Documenta Mathematica for various kind suggestions. During this research, JP was partially supported by the National Research Foundation of Korea (NRF) grants No. 2013042157 and No. 2015R1A2A2A01004120, Korea Institute for Advanced Study (KIAS) grant, all funded by the Korean government (MSIP), and TJ Park Junior Faculty Fellowship funded by POSCO TJ Park Foundation.

References

[1] F. Binda, S. Saito, *Relative cycles with moduli and regulator maps*, arXiv:1412.0385, (2014).

[2] S. Bloch, *Algebraic K-theory and crystalline cohomology*, Publ. Math. de l'Inst. Haut. Etud. Sci., 47, (1977), 187–268.

[3] S. Bloch, *Algebraic cycles and higher K-theory*, Adv. Math., 61, (1986), 267–304.

[4] S. Bloch, *The moving lemma for higher Chow groups*, J. Algebraic Geom., 3, (1994), no. 3, 537–568.

[5] S. Bloch, H. Esnault, *The additive dilogarithm*, Doc. Math., Extra Vol., (2003), 131–155.

[6] W. Fulton, *Intersection theory*, Second Edition, Ergebnisse der Mathematik und ihrer Grenzgebiete 3, Folge. Modern Surveys in Mathematics, 2, Springer-Verlag, Berlin, 1998.

[7] R. Hartshorne, *Algebraic Geometry*, Graduate Texts in Math., 52, Springer-Verlag, New York., 1977.

[8] L. Hesselholt, I. Madsen, *On the K-theory of nilpotent endomorphisms*, in Homotopy methods in Algebraic Topology (Bounder, CO, 1999), edited by Greenlees, J. P. C. et al., 127–140. Contempo. Math. 271, Providence, RI, American Math. Soc. 2001.

[9] L. Hesselholt, *K-theory of truncated polynomial algebras*, in Handbook of K-Theory, Volume 1, Springer-Verlag, Berlin, 2005, pp. 71–110.

[10] L. Hesselholt, *On the K-theory of coordinate axes in the plane*, Nagoya Math. J., 185, (2007), 93–109.

[11] L. Hesselholt, *The big de Rham-Witt complex*, Acta Math., 214, (2015), 135–207.

[12] L. Illusie, *Complexe de de Rham-Witt et cohomologie cristalline*, Ann Sci. de l'Ecole Norm. Sup. 4^e série, 12, (1979), 501–661.

[13] J. F. Jardine, *A closed model structure for differential graded algebras*, in Cyclic cohomology and non-commutative geometry (Waterloo, ON, 1995) 55–58, Fields Inst. Commun., 17, Amer. Math. Soc., Providence, RI, 1997.

[14] W. Kai, *A moving lemma for algebraic cycles with modulus and contravariance*, arXiv:1507.07619v1, (2015).

[15] M. Kerz, S. Saito, *Chow group of 0-cycles with modulus and higher dimensional class field theory*, arXiv:1304.4400, to appear in Duke Math. J.

[16] A. Krishna, M. Levine, *Additive higher Chow groups of schemes*, J. Reine Angew. Math., 619, (2008) 75–140.

[17] A. Krishna, J. Park, *Moving lemma for additive higher Chow groups*, Algebra Number Theory, 6, (2012), no. 2, 293–326.

[18] A. Krishna, J. Park, *Mixed motives over* $k[t]/(t^{m+1})$, J. Inst. Math. Jussieu, 11, (2012), no. 1, 611–657.

[19] A. Krishna, J. Park, *DGA-structure on additive higher Chow groups*, Int. Math. Res. Not. 2015, no. 1, 1–54.

[20] A. Krishna, J. Park, *Semitopologization in motivic homotopy theory and applications*, Alg. Geom. Top., 15, (2015), no. 2, 823–861.

[21] A. Krishna, J. Park, *A module structure and a vanishing theorem for cycles with modulus*, arXiv:1412.7396v2, (2015).

[22] A. Krishna, J. Park, *Algebraic cycles and crystalline cohomology*, arXiv:1504.08181, (2015).

[23] M. Levine, *Bloch's higher Chow groups revisited*, K-theory (Strasbourg, 1992). Astérisque, 226, (1994), 235–320.

[24] M. Levine, *Smooth motives*, in Motives and algebraic cycles, 175-231, Fields Inst. Commun., 56, Amer. Math. Soc., Providence, RI, 2009.

[25] Q. Liu, *Algebraic Geometry and Arithmetic Curves*, Oxford Graduate Texts in Mathematics, Vol. 6. Oxford Univ. Press, London (2006).

[26] J. Park, *Algebraic cycles and additive dilogarithm*, Int. Math. Res. Not., 2007, no. 18, Article ID rnm067.

[27] J. Park, *Regulators on additive higher Chow groups*, Amer. J. Math., 131, (2009), no. 1, 257–276.

[28] K. Rülling, *The generalized de Rham-Witt complex over a field is a complex of zero-cycles*, J. Algebraic Geom., 16, (2007), no. 1, 109–169.

Amalendu Krishna
School of Mathematics
Tata Institute of
Fundamental Research
1 Homi Bhabha Road
Colaba, Mumbai, India
amal@math.tifr.res.in

Jinhyun Park
Department of Mathematical Sciences
KAIST 291 Daehak-ro Yuseong-gu
Daejeon 34141
Republic of Korea (South)
jinhyun@mathsci.kaist.ac.kr
jinhyun@kaist.edu

Tame Class Field Theory for Singular Varieties over Algebraically Closed Fields

Thomas Geisser[1] and Alexander Schmidt[2]

Received: May 25, 2015
Revised: November 25, 2015

Communicated by Stephen Lichtenbaum

Abstract. Let X be a separated scheme of finite type over an algebraically closed field k and let m be a natural number. By an explicit geometric construction using torsors we construct a pairing between the first mod m Suslin homology and the first mod m tame étale cohomology of X. We show that the induced homomorphism from the mod m Suslin homology to the abelianized tame fundamental group of X mod m is surjective. It is an isomorphism of finite abelian groups if $(m, \mathrm{char}(k)) = 1$, and for general m if resolution of singularities holds over k.

2010 Mathematics Subject Classification: 14F35, 14F43, 14C25
Keywords and Phrases: Suslin homology, higher dimensional class field theory, tame fundamental group

1 Introduction

Let X be a (possibly singular) separated scheme of finite type over an algebraically closed field k of characteristic $p \geq 0$ and let m be a natural number. We construct a pairing between the first mod m algebraic singular homology $H_1^S(X, \mathbb{Z}/m\mathbb{Z})$ and the first mod m tame étale cohomology group $H_t^1(X, \mathbb{Z}/m\mathbb{Z})$. For $\pi_1^{t,ab}(X) = H_t^1(X, \mathbb{Q}/\mathbb{Z})^\vee$ we prove the following analogue of Hurewicz's theorem in algebraic topology:

[1]Supported by JSPS Grant-in-Aid (B) 23340004
[2]Supported by DFG-Forschergruppe FOR 1920

THEOREM 1.1. *The induced homomorphism*

$$rec_X : H_1^S(X, \mathbb{Z}/m\mathbb{Z}) \longrightarrow \pi_1^{t,ab}(X)/m$$

is surjective. It is an isomorphism of finite abelian groups if $(m, p) = 1$*, and for general* m *if resolution of singularities holds for schemes of dimension* $\leq \dim X + 1$ *over* k*.*

For $p \nmid m$, the groups $H_1^S(X, \mathbb{Z}/m\mathbb{Z})$ and $\pi_1^{t,ab}(X)/m$ are known to be isomorphic by the work of Suslin and Voevodsky [SV1]. Theorem 1.1 above provides an explicit isomorphism which extends to the case $p \mid m$ (under resolution of singularities). Moreover, in the last section we show that for $p \nmid m$ our isomorphism coincides with the one constructed in [SV1].

The motivation for constructing our pairing between the groups $H_1^S(X, \mathbb{Z}/m\mathbb{Z})$ and $H_t^1(X, \mathbb{Z}/m\mathbb{Z})$ comes from topology: For a locally contractible Hausdorff space X and a natural number m, the canonical duality pairing

$$\langle \cdot, \cdot \rangle : H_1^{\rm sing}(X, \mathbb{Z}/m\mathbb{Z}) \times H^1(X, \mathbb{Z}/m\mathbb{Z}) \longrightarrow \mathbb{Z}/m\mathbb{Z},$$

between singular homology and sheaf cohomology with mod m coefficients can be given explicitly in the following way: represent $b \in H^1(X, \mathbb{Z}/m\mathbb{Z})$ by a $\mathbb{Z}/m\mathbb{Z}$-torsor $\mathcal{T} \to X$ and $a \in H_1^{\rm sing}(X, \mathbb{Z}/m\mathbb{Z})$ by a 1-cycle α in the singular complex of X. Then

$$\langle a, b \rangle = \Phi_{par}^{-1} \circ \Phi_{taut} \in \mathbb{Z}/m\mathbb{Z}, \text{ where } \Phi_{taut}, \Phi_{par} : \alpha^*(\mathcal{T})|_0 \xrightarrow{\sim} \alpha^*(\mathcal{T})|_1$$

are the isomorphisms between the fibres over 0 and 1 of the pull-back torsor $\alpha^*(\mathcal{T}) \to \Delta^1 = [0,1]$ given tautologically ($0^*\alpha = 1^*\alpha$) and by parallel transport (every $\mathbb{Z}/m\mathbb{Z}$-torsor on $[0,1]$ is trivial).

For a variety X, the pairing between $H_1^S(X, \mathbb{Z}/m\mathbb{Z})$ and $H_t^1(X, \mathbb{Z}/m\mathbb{Z})$ inducing the homomorphism rec_X of our Main Theorem 1.1 will be constructed in the same way. However, 1-cycles in the algebraic singular complex are not linear combinations of morphisms but finite correspondences from Δ^1 to X. In order to mimic the above construction, we thus have to define the pull-back of a torsor along a finite correspondence, which requires the construction of the push-forward torsor along a finite surjective morphism.

To prove Theorem 1.1, we first consider the case of a smooth curve C. If $\mathcal{A}$ is the Albanese variety of C, then we have isomorphisms

$$H_1^S(C, \mathbb{Z}/m\mathbb{Z}) \underset{\sim}{\xrightarrow{\delta}} {}_mH_0^S(C, \mathbb{Z}) \cong {}_m\mathcal{A}(k). \tag{1}$$

The first isomorphism follows from the coefficient sequence together with the divisibility of $H_1^S(C, \mathbb{Z})$, and the second from the Abel-Jacobi theorem. On the other hand,

$$\mathrm{Hom}({}_mA(k), \mathbb{Z}/m\mathbb{Z}) \underset{\sim}{\xrightarrow{\tau}} H_t^1(C, \mathbb{Z}/m\mathbb{Z}). \tag{2}$$

This follows because the maximal étale subcovering $\widetilde{\mathcal{A}} \to \mathcal{A}$ of the m-multiplication map $\mathcal{A} \stackrel{m}{\to} \mathcal{A}$ is the quotient of $\mathcal{A}$ by the connected component of the finite group scheme ${}_m\mathcal{A}$, and the maximal abelian tame étale covering of C with Galois group annihilated by m is $\widetilde{C} := C \times_{\mathcal{A}} \widetilde{\mathcal{A}}$. The heart of the proof of Theorem 1.1 for smooth curves is to show that under the above identifications, our pairing agrees with the evaluation map.

We then show surjectivity of rec_X for general X by reducing to the case of smooth curves. Finally, we use duality theorems to show that both sides of rec_X have the same order: For the p-primary part, we use resolution of singularities to reduce to the smooth projective case considered in [Ge3]. For $(m, \mathrm{char}(k)) = 1$, Suslin and Voevodsky [SV1] construct an isomorphism

$$\alpha_X : H^1_{\mathrm{et}}(X, \mathbb{Z}/m\mathbb{Z}) \stackrel{\sim}{\to} H^1_S(X, \mathbb{Z}/m\mathbb{Z}).$$

Hence the source and the target of rec_X have the same order and therefore rec_X is an isomorphism. In Section 7 we show that rec_X is dual to the map α_X. Thus, for $\mathrm{char}(k) \nmid m$, our construction gives an explicit description of the Suslin-Voevodsky isomorphism α_X, which zig-zags through Ext-groups in various categories and is difficult to understand.

The authors thank Takeshi Saito and Changlong Zhong for discussions during the early stages of the project. It is a pleasure to thank Johannes Anschütz whose comments on an earlier version of this paper led to a substantial simplification of the proof of Theorem 4.1. Finally, we want to thank the referee for his helpful comments.

2 TORSORS AND FINITE CORRESPONDENCES

All occurring schemes in this section are separated schemes of finite type over a field k. For any abelian group A and a finite surjective morphism $\pi : Z \to X$ with Z integral and X normal, connected, we have transfer maps

$$\pi_* : H^i_{\mathrm{et}}(Z, A) \to H^i_{\mathrm{et}}(X, A)$$

for all $i \geq 0$ (see [MVW], 6.11, 6.21). The group $H^1_{\mathrm{et}}(Z, A)$ classifies isomorphism classes of étale A-torsors (i.e., principal homogeneous spaces) over the scheme Z. We are going to construct a functor

$$\pi_* : \mathcal{PHS}(Z, A) \longrightarrow \mathcal{PHS}(X, A)$$

from the category of étale A-torsors on Z to the category of étale A-torsors on X, which induces the transfer map $\pi_* : H^1_{\mathrm{et}}(Z, A) \to H^1_{\mathrm{et}}(X, A)$ above on isomorphism classes.

We recall how to add and subtract torsors. For an abelian group A and A-torsors $\mathcal{T}_1$, $\mathcal{T}_2$ on a scheme Y, define

$$\mathcal{T}_1 + \mathcal{T}_2$$

to be the quotient scheme of $\mathcal{T}_1 \times_Y \mathcal{T}_2$ by the action of A given by $(t_1, t_2) + a = (t_1 + a, t_2 - a)$. It carries the structure of an A-torsor by setting

$$\overline{(t_1, t_2)} + a := \overline{(t_1 + a, t_2)} \quad (= \overline{(t_1, t_2 + a)}).$$

The functor

$$+ : \mathcal{PHS}(Y, A) \times \mathcal{PHS}(Y, A) \longrightarrow \mathcal{PHS}(Y, A), \quad (\mathcal{T}_1, \mathcal{T}_2) \longmapsto \mathcal{T}_1 + \mathcal{T}_2,$$

lifts the addition in $H^1_{\mathrm{et}}(Y, A)$ to torsors (cf. [Mi], III, Rem. 4.8 (b)). Note that "$+$" is associative and commutative up to natural functor isomorphisms. In particular, we can multiply a torsor by any natural number m, putting $m \cdot \mathcal{T} = \mathcal{T} + \cdots + \mathcal{T}$ (m times). If $mA = 0$, then we have a natural isomorphism of torsors

$$m \cdot \mathcal{T} \xrightarrow{\sim} Y \times A, \quad \overline{(t_1, \ldots, t_m)} \mapsto (t_2 - t_1) + \cdots + (t_m - t_1) \in A, \tag{3}$$

where $Y \times A$ is the trivial A-torsor on Y representing the constant sheaf $\underline{A}$ over Y. Here $t_i - t_j$ denotes the unique element $a \in A$ with $t_i = t_j + a$.

Furthermore, given a torsor $\mathcal{T}$, define $(-\mathcal{T})$ to be the torsor which is isomorphic to $\mathcal{T}$ as a scheme and on which $a \in A$ acts as $-a$. This yields a functor

$$(-1) : \mathcal{PHS}(Y, A) \longrightarrow \mathcal{PHS}(Y, A), \quad \mathcal{T} \longmapsto (-\mathcal{T}),$$

which lifts multiplication by (-1) from $H^1_{\mathrm{et}}(Y, A)$ to an endofunctor of $\mathcal{PHS}(Y, A)$. We have a natural isomorphism of torsors

$$\mathcal{T} + (-\mathcal{T}) \xrightarrow{\sim} Y \times A, \quad \overline{(t_1, t_2)} \mapsto t_1 - t_2 \in A. \tag{4}$$

Now let $\pi : Z \to X$ be finite and surjective, Z integral, X normal, connected, and let $\mathcal{T}$ be an A-torsor on Z. For every point $x \in X$, the base change $Z \times_X X_x^{sh}$ is a product of strictly henselian local schemes. Therefore we find an étale cover $(U_i \to X)_{i \in I}$ of X such that $\mathcal{T}$ trivializes over the pull-back étale cover $(\pi^{-1}(U_i) \to Z)_{i \in I}$ of Z.
Next choose a pseudo-Galois covering $\widetilde{\pi} : \widetilde{Z} \to X$ dominating $Z \to X$. Recall that this means that $k(\widetilde{Z})|k(X)$ is a normal field extension and that the natural map $\mathrm{Aut}_X(\widetilde{Z}) \to \mathrm{Aut}_{k(X)}(k(\widetilde{Z}))$ is bijective (cf. [SV1], Lemma 5.6). Let $\pi_{in} : X_{in} \to X$ be the quotient scheme $\widetilde{Z}/G$, where $G = \mathrm{Aut}_X(\widetilde{Z})$. Then X_{in} is the normalization of X in the maximal purely inseparable subextension $k(X)^{in}/k(X)$ of $k(\widetilde{Z})/k(X)$. Consider the object

$$\widetilde{\mathcal{T}} := \sum_{\varphi \in \mathrm{Mor}_X(\widetilde{Z}, Z)} \varphi^*(\mathcal{T}) \ \in \ \mathcal{PHS}(\widetilde{Z}, A),$$

which is defined up to unique isomorphism. Starting from any trivialization of $\mathcal{T}$ over $(\pi^{-1}(U_i) \to Z)_{i \in I}$, we obtain a trivialization of the restriction of $\widetilde{\mathcal{T}}$ to $(\widetilde{\pi}^{-1}(U_i) \to \widetilde{Z})_{i \in I}$ of the form

$$\widetilde{\mathcal{T}}|_{\widetilde{\pi}^{-1}(U_i)} \cong \widetilde{\pi}^{-1}(U_i) \times A,$$

where $G = \mathrm{Aut}_X(\widetilde{Z})$ acts on the right hand side in the canonical way on $\widetilde{\pi}^{-1}(U_i)$ and trivially on A. Therefore the quotient scheme $\widetilde{\mathcal{T}}/G$ is an A-torsor on $\widetilde{Z}/G = X_{in}$ in a natural way. Since $X_{in} \to X$ is a topological isomorphism, $\widetilde{\mathcal{T}}/G$ comes by base change from a unique A-torsor $\mathcal{T}'$ on X.

DEFINITION 2.1. *The push-forward A-torsor $\pi_*(\mathcal{T})$ on X is defined by*

$$\pi_*(\mathcal{T}) = [k(Z) : k(X)]_{in} \cdot \mathcal{T}'.$$

The assignment $\mathcal{T} \mapsto \pi_*(\mathcal{T})$ defines a functor

$$\pi_* : \mathcal{PHS}(Z, A) \longrightarrow \mathcal{PHS}(X, A).$$

The functor π_* is additive in the sense that it commutes with the functors "$+$" and "(-1)" up to a natural functor isomorphism.

Let $\mathcal{T} \in \mathcal{PHS}(Z, A)$ and assume that there exists a section $s : Z \to \mathcal{T}$ to the projection $\mathcal{T} \to Z$ (so $\mathcal{T}$ is trivial and s gives a trivialization). Let again $\pi : Z \to X$ be finite and surjective, Z integral, X normal, connected. Then

$$\widetilde{\mathcal{T}} := \sum_{\varphi \in \mathrm{Mor}_X(\widetilde{Z}, Z)} \varphi^*(\mathcal{T}) \in \mathcal{PHS}(\widetilde{Z}, A)$$

has the canonical section $\sum_{\varphi \in \mathrm{Mor}_X(\widetilde{Z}, Z)} \varphi^*(s)$ over $\widetilde{Z}$. It descends to a section of $\mathcal{T}/G$ over $\widetilde{Z}/G = X_{in}$. Descending to X and multiplying by $[k(X_{in}) : k(X)]$, we obtain a section

$$\pi_*(s) : X \to \pi_*(\mathcal{T}).$$

In other words, we obtain a map

$$\pi_* : \Gamma(Z, \mathcal{T}) \longrightarrow \Gamma(X, \pi_*(\mathcal{T}));$$

hence every trivialization of $\mathcal{T}$ gives a trivialization of $\pi_*(\mathcal{T})$ in a natural way.

In order to see that π_* induces the transfer map $\pi_* : H^1_{\mathrm{et}}(Z, A) \to H^1_{\mathrm{et}}(X, A)$ after passing to isomorphism classes, we formulate the construction of π_* on the level of Čech 1-cocycles. As explained above, we find an étale cover $(U_i \to X)_{i \in I}$ such that $\mathcal{T}$ trivializes over the étale cover $(\pi^{-1}(U_i) \to Z)_{i \in I}$ of Z. We fix a trivialization and obtain a Čech 1-cocycle

$$a = (a_{ij} \in \Gamma(\pi^{-1}(U_i \times_X U_j), A))$$

over $(\pi^{-1}(U_i) \to Z)_{i \in I}$ which defines $\mathcal{T}$. As before choose a pseudo-Galois covering $\widetilde{\pi} : \widetilde{Z} \to X$ dominating $Z \to X$. Now for all i, j consider the element

$$\sum_{\varphi \in \mathrm{Mor}_X(\widetilde{Z}, Z)} \varphi^*(a_{ij}) \in \Gamma(\widetilde{\pi}^{-1}(U_i \times_X U_j), A)$$

which, by Galois invariance, lies in

$$\Gamma(\pi_{in}^{-1}(U_i \times_X U_j), A) = \Gamma(U_i \times_X U_j, A).$$

The Čech 1-cocycle given by

$$[k(Z) : k(X)]_{in} \cdot \Big(\sum_{\varphi \in \mathrm{Mor}_X(\widetilde{Z}, Z)} \varphi^*(a_{ij}) \Big) \in \Gamma(U_i \times_X U_j, A).$$

now defines a trivialization of $\pi_*(\mathcal{T})$ over $(U_i \to X)_{i \in I}$. Since the transfer map on étale cohomology is defined on Čech cocycles in exactly this way (see [MVW], 6.11, 6.21), we obtain

LEMMA 2.2. *Passing to isomorphism classes, the functor* $\pi_* : \mathcal{PHS}(Z, A) \to \mathcal{PHS}(X, A)$ *constructed above induces the transfer homomorphism*

$$\pi_* : H^1_{\mathrm{et}}(Z, A) \to H^1_{\mathrm{et}}(X, A).$$

If any finite subset of closed points of X is contained in an affine open, then symmetric powers exist, and another description of the push-forward for torsors is the following: Associated with the finite morphism $\pi : Z \to X$ of degree d, there is a section $s_\pi : X \to \mathrm{Sym}^d(Z/X)$ to the natural projection $\mathrm{Sym}^d(Z/X) \to X$ (see ([SV1], p. 81). We denote the composite of s_π with $pr : \mathrm{Sym}^d(Z/X) \to \mathrm{Sym}^d(Z)$ by S_π. Defining $f : \widetilde{Z} \to \mathrm{Sym}^d(Z/X)$ by repeating each element in $\mathrm{Mor}_X(\widetilde{Z}, Z)$ exactly $[k(Z) : k(X)]_{in}$-times, the diagram

$$\begin{tikzcd} \widetilde{Z} \arrow[r, "f"] \arrow[dr, "\widetilde{\pi}"'] & \mathrm{Sym}^d(Z/X) \arrow[r, "pr"] & \mathrm{Sym}^d(Z) \\ & X \arrow[u, "s_\pi"] \arrow[ur, "S_\pi"'] & \end{tikzcd}$$

commutes. For an A-torsor $\mathcal{T} \to Z$, the d-fold self-product $\mathcal{T} \times_k \cdots \times_k \mathcal{T}$ is an A^d-torsor over the d-fold self-product of Z in a natural way. Taking the quotient by the A^{d-1}-action

$$(a_1, \ldots, a_{d-1})(t_1, \ldots, t_d) = (t_1 + a_1, t_2 - a_1 + a_2, t_3 - a_2 + a_3, \ldots, t_d - a_{d-1}),$$

we obtain an A-torsor over Z^d. Dividing out the by the action of the symmetric group S_d, we obtain an A-torsor over $\mathrm{Sym}^d(Z)$ and denote it by $\mathrm{Sym}^d_A(\mathcal{T})$. We obtain natural isomorphisms in $\mathcal{PHS}(\widetilde{Z}, A)$:

$$\begin{aligned} [k(Z) : k(X)]_{in} \cdot \sum_{\varphi \in \mathrm{Mor}_X(\widetilde{Z}, Z)} \varphi^*(\mathcal{T}) &\cong (pr \circ f)^* \,\mathrm{Sym}^d_A(\mathcal{T}) \\ &\cong \widetilde{\pi}^* \circ (pr \circ s_\pi)^* \,\mathrm{Sym}^d_A(\mathcal{T}). \end{aligned}$$

By our construction of $\pi_*(\mathcal{T})$ we obtain

LEMMA 2.3. *We have a natural isomorphism in $\mathcal{PHS}(X, A)$:*

$$\pi_*(\mathcal{T}) \cong S_\pi^*(\mathrm{Sym}^d(\mathcal{T})),$$

where $S_\pi = pr \circ s_\pi : X \to \mathrm{Sym}^d(Z)$.

Assume now that X is regular and Y arbitrary. The group of finite correspondences $\mathrm{Cor}(X, Y)$ is defined as the free abelian group on the set of integral subschemes $Z \subset X \times Y$ which project finitely and surjectively to a connected component of X. For such a Z, we define $p_{[Z\to X]\,*} : \mathcal{PHS}(Z, A) \to \mathcal{PHS}(X, A)$ by extending (if X is not connected) the push-forward torsor defined above in a trivial way to those connected components of X which are not dominated by Z. We consider the functor

$$[Z]^* = p_{[Z\to X]\,*} \circ p^*_{[Z\to Y]} : \mathcal{PHS}(Y, A) \longrightarrow \mathcal{PHS}(X, A).$$

Using the operations "$+$" and "(-1)" we extend this construction to arbitrary finite correspondences.

DEFINITION 2.4. *Let X be regular, Y arbitrary and $\alpha = \sum n_i Z_i \in \mathrm{Cor}(X, Y)$ a finite correspondence. Then*

$$\alpha^* : \mathcal{PHS}(Y, A) \longrightarrow \mathcal{PHS}(X, A)$$

is defined by setting

$$\alpha^*(\mathcal{T}) := \sum n_i [Z_i]^*(\mathcal{T}).$$

Using the isomorphism (4) above, we immediately obtain

LEMMA 2.5. *For $\alpha_1, \alpha_2 \in Cor(X, Y)$ and $\mathcal{T}_1, \mathcal{T}_2 \in \mathcal{PHS}(Y, A)$, $n_1, n_2 \in \mathbb{Z}$, we have a natural isomorphism*

$$(\alpha_1 + \alpha_2)^*(n_1\mathcal{T}_1 + n_2\mathcal{T}_2) \cong n_1\alpha_1^*(\mathcal{T}_1) + n_1\alpha_2^*(\mathcal{T}_1) + n_2\alpha_1^*(\mathcal{T}_2) + n_2\alpha_2^*(\mathcal{T}_2).$$

If X and Y are regular and Z is arbitrary, we have a natural composition law

$$\mathrm{Cor}(X, Y) \times \mathrm{Cor}(Y, Z) \longrightarrow \mathrm{Cor}(X, Z),\ (\alpha, \beta) \mapsto \beta \circ \alpha,$$

(see [MVW], Lecture 1). A straightforward but lengthy computation unfolding the definitions shows

PROPOSITION 2.6. *Let X and Y be regular and Z arbitrary. Let $\alpha \in \mathrm{Cor}(X,Y)$ and $\beta \in \mathrm{Cor}(Y,Z)$. Then, for any $\mathcal{T} \in \mathcal{PHS}(Z,A)$, we have a canonical isomorphism*

$$\alpha^*(\beta^*(\mathcal{T})) \cong (\beta \circ \alpha)^*(\mathcal{T}).$$

Finally, assume that $mA = 0$ for some natural number m. Then (using the isomorphism (3) above), we have for any $\alpha, \beta \in Cor(X,Y)$, $\mathcal{T} \in \mathcal{PHS}(Y,A)$, a natural isomorphism

$$(\alpha + m\beta)^*(\mathcal{T}) \cong \alpha^*(\mathcal{T}).$$

Therefore, we have an A-torsor

$$\bar{\alpha}^*(\mathcal{T}) \in \mathcal{PHS}(X,A)$$

given up to unique isomorphism for any $\bar{\alpha} \in \mathrm{Cor}(X,Y) \otimes \mathbb{Z}/m\mathbb{Z}$. In other words, we obtain the

LEMMA 2.7. *Assume that $mA = 0$, and let $\alpha, \beta \in \mathrm{Cor}(X,Y)$ have the same image in $\mathrm{Cor}(X,Y) \otimes \mathbb{Z}/m\mathbb{Z}$. Then there is a natural isomorphism of functors*

$$\alpha^* \cong \beta^* : \mathcal{PHS}(Y,A) \to \mathcal{PHS}(X,A).$$

For a regular connected curve C we consider the subgroup $H^1_t(C,A) \subseteq H^1_{\mathrm{et}}(C,A)$ of tame cohomology classes (corresponding to those continuous homomorphisms $\pi_1^{\mathrm{et}}(C) \to A$ which factor through the tame fundamental group $\pi_1^t(\bar{C}, \bar{C} - C)$, where $\bar{C}$ is the unique regular compactification of C).
For a general scheme X over k we call a cohomology class in $a \in H^1_{\mathrm{et}}(X,A)$ curve-tame (or just tame) if for any morphism $f : C \to X$ with C a regular curve, we have $f^*(a) \in H^1_t(C,A)$. The tame cohomology classes form a subgroup

$$H^1_t(X,A) \subseteq H^1_{\mathrm{et}}(X,A).$$

The groups coincide if X is proper or if $p = 0$ or if $p > 0$ and A is p-torsion free, where p is the characteristic of the base field k.

DEFINITION 2.8. *We call an étale A-torsor $\mathcal{T}$ on X* tame *if its isomorphism class lies in $H^1_t(X,A) \subseteq H^1_{\mathrm{et}}(X,A)$.*

LEMMA 2.9. *Let Z be integral, X normal, connected, $\pi : Z \to X$ finite, surjective and $f : Z \to Y$ any morphism. Let $\mathcal{T}$ be a tame torsor on Y. Then $\pi_*(f^*(\mathcal{T}))$ is a tame torsor on X.*

Proof. By definition, f^* preserves curve-tameness. So we may assume $Z = Y$, $f = \mathrm{id}$. Again by the definition of curve-tameness and using Proposition 2.6, we may reduce to the case that X is a regular curve. Since étale cohomology commutes with direct limits of coefficients, we may assume that A is a finitely generated abelian group. Furthermore, we may assume that $\mathrm{char}(k) = p > 0$ and $A = \mathbb{Z}/p^r\mathbb{Z}$, $r \geq 1$. Let $\bar{Z}$ be the canonical compactification of Z, i.e., the unique proper curve over k which contains Z as a dense open subscheme and such that all points of $\bar{Z} \smallsetminus Z$ are regular points of $\bar{Z}$. By the definition of tame coverings of curves, $\mathcal{T}$ extends to a $\mathbb{Z}/p^r\mathbb{Z}$-torsor on $\bar{Z}$. Hence also $\pi_*(\mathcal{T})$ extends to the canonical compactification $\bar{X}$ of X and so is tame. □

PROPOSITION 2.10. *Let $\bar{X}$ be a proper and regular scheme over k and let $X \subset \bar{X}$ be a dense open subscheme. Let $p = \mathrm{char}(k) > 0$. Then for any $r \geq 1$ the natural inclusion*
$$H^1_{\mathrm{et}}(\bar{X}, \mathbb{Z}/p^r\mathbb{Z}) \hookrightarrow H^1_{\mathrm{et}}(X, \mathbb{Z}/p^r\mathbb{Z})$$
induces an isomorphism
$$H^1_{\mathrm{et}}(\bar{X}, \mathbb{Z}/p^r\mathbb{Z}) = H^1_t(\bar{X}, \mathbb{Z}/p^r\mathbb{Z}) \xrightarrow{\sim} H^1_t(X, \mathbb{Z}/p^r\mathbb{Z}) \subseteq H^1_{\mathrm{et}}(X, \mathbb{Z}/p^r\mathbb{Z}).$$

Proof. Let $\mathcal{T}_0$ be any connected component of a tame $\mathbb{Z}/p^r\mathbb{Z}$-torsor $\mathcal{T}$ on X. Then the morphism $\mathcal{T}_0 \to X$ is curve-tame in the sense of [KS], §4, and $\mathcal{T}_0$ is the normalization of X in the abelian field extension of p-power degree $k(\mathcal{T}_0)/k(X)$. By [KS], Thm. 5.4. (b), $\mathcal{T}_0 \to X$ is numerically tamely ramified along $\bar{X} \smallsetminus X$. This means that the inertia groups in $\mathrm{Gal}(k(\mathcal{T}_0)/k(X))$ of all points $\bar{x} \in \bar{X} \smallsetminus X$ are of order prime to p, hence trivial. Therefore $\mathcal{T}_0$, and thus $\mathcal{T}$ extends to $\bar{X}$. □

COROLLARY 2.11. *Let $\Delta^n = Spec(k[T_0, \ldots, T_n]/\sum T_i = 1)$ be the n-dimensional standard simplex over k and let A be an abelian group. Then*
$$H^1_t(\Delta^n, A) \cong H^1_{\mathrm{et}}(k, A).$$
In particular, $H^1_t(\Delta^n, A) = 0$ if k is separably closed.

Proof. Since tame cohomology commutes with direct limits of coefficients, and since $H^1_{\mathrm{et}}(\Delta^n, \mathbb{Z}) = 0$, we may assume that $A \cong \mathbb{Z}/m\mathbb{Z}$ for some $m \geq 1$. If $p \nmid m$, we obtain:
$$H^1_t(\Delta^n, \mathbb{Z}/m\mathbb{Z}) \cong H^1_{\mathrm{et}}(\mathbb{A}^n, \mathbb{Z}/m\mathbb{Z}) \cong H^1_{\mathrm{et}}(k, \mathbb{Z}/m\mathbb{Z}).$$
If $p = \mathrm{char}(k) > 0$ and $m = p^r$, $r \geq 1$, Proposition 2.10 yields
$$H^1_t(\Delta^n, \mathbb{Z}/p^r\mathbb{Z}) \cong H^1_t(\mathbb{A}^n, \mathbb{Z}/p^r\mathbb{Z}) \xleftarrow{\sim} H^1_t(\mathbb{P}^n, \mathbb{Z}/p^r\mathbb{Z}) = H^1_{\mathrm{et}}(\mathbb{P}^n, \mathbb{Z}/p^r\mathbb{Z}).$$
Finally note that $H^1_{\mathrm{et}}(\mathbb{P}^n, \mathbb{Z}/p^r\mathbb{Z}) \cong H^1_{\mathrm{et}}(k, \mathbb{Z}/p^r\mathbb{Z})$. □

In the following, let k be an algebraically closed field of characteristic $p \geq 0$ and let X be a separated scheme of finite type over k. Let $H_i^S(X, \mathbb{Z}/m\mathbb{Z})$ denote the mod-m Suslin homology, i.e., the i-th homology group of the complex

$$\mathrm{Cor}(\Delta^\bullet, X) \otimes \mathbb{Z}/m\mathbb{Z}.$$

Let A be an abelian group with $mA = 0$. We are going to construct a pairing

$$H_1^S(X, \mathbb{Z}/m\mathbb{Z}) \times H_t^1(X, A) \longrightarrow A$$

as follows: let $\mathcal{T} \to X$ be a tame A-torsor representing a class in $H_t^1(X, A)$ and let $\alpha \in \mathrm{Cor}(\Delta^1, X)$ be a finite correspondence representing a 1-cocycle in the mod-m Suslin complex. Then

$$\alpha^*(\mathcal{T})$$

is a torsor over Δ^1. Since α is a cocycle modulo m, $(0^* - 1^*)(\alpha)$ is of the form $m \cdot z$ for some $z \in \mathrm{Cor}(\Delta^0, X) = \mathbb{Z}^{(X(k))}$. We therefore obtain a canonical identification

$$\Phi_{taut} : 0^*(\alpha^*(\mathcal{T})) \xrightarrow{\sim} 1^*(\alpha^*(\mathcal{T}))$$

of A-torsors over $\Delta^0 = \mathit{Spec}(k)$. Furthermore, by Corollary 2.11, the tame torsor $\alpha^*(\mathcal{T})$ on Δ^1 is trivial, hence a disjoint union of copies of Δ^1. By parallel transport, we obtain another identification

$$\Phi_{par} : 0^*(\alpha^*(\mathcal{T})) \xrightarrow{\sim} 1^*(\alpha^*(\mathcal{T})).$$

Hence there is a unique $\gamma(\alpha, \mathcal{T}) \in A$ such that

$$\Phi_{par} = (\text{translation by } \gamma(\alpha, \mathcal{T})) \circ \Phi_{taut}.$$

PROPOSITION 2.12. *The element $\gamma(\alpha, \mathcal{T}) \in A$ only depends on the class of $\mathcal{T}$ in $H_t^1(X, A)$ and on the class of α in $H_1^S(X, \mathbb{Z}/m\mathbb{Z})$. We obtain a bilinear pairing*

$$\langle \cdot, \cdot \rangle : H_1^S(X, \mathbb{Z}/m\mathbb{Z}) \times H_t^1(X, A) \longrightarrow A.$$

Proof. Replacing $\mathcal{T}$ by another torsor isomorphic to $\mathcal{T}$ does not change anything. The nontrivial statement is that $\langle \alpha, \mathcal{T} \rangle$ only depends on the class of α in $H_1^S(X, \mathbb{Z}/m\mathbb{Z})$. For $\beta \in \mathrm{Cor}(\Delta^1, X)$, we have

$$\langle \alpha + m\beta, \mathcal{T} \rangle = \langle \alpha, \mathcal{T} \rangle + m\langle \beta, \mathcal{T} \rangle = \langle \alpha, \mathcal{T} \rangle.$$

It therefore remains to show that

$$\langle \partial^*(\Phi), \mathcal{T} \rangle = 0,$$

for all $\Phi \in \mathrm{Cor}(\Delta^2, X)$, where $\partial_i : \Delta^1 \to \Delta^2$, $i = 0, 1, 2$, are the face maps and $\partial^*(\Phi) = \Phi \circ \partial_0 - \Phi \circ \partial_1 + \Phi \circ \partial_2$. Considering $\partial = \partial_0 - \partial_1 + \partial_2$ as a finite correspondence from Δ^1 to Δ^2, it represents a cocycle in the singular complex $\mathrm{Cor}(\Delta^\bullet, \Delta^2)$. Proposition 2.6 implies that

$$\langle \partial^*(\Phi), \mathcal{T}\rangle = \langle \Phi \circ \partial, \mathcal{T}\rangle = \langle \partial, \Phi^*(\mathcal{T})\rangle.$$

By Corollary 2.11, the tame torsor $\Phi^*(\mathcal{T})$ is trivial on Δ^2. Hence $\langle \partial, \Phi^*(\mathcal{T})\rangle = 0$. □

In the following, we use the notation $\pi_1^{t,ab}(X) := H_t^1(X, \mathbb{Q}/\mathbb{Z})^*$. If X is connected, then $\pi_1^{t,ab}(X)$ is the abelianized (curve-)tame fundamental group of X, see [KS], §4.

DEFINITION 2.13. *For $m \geq 1$ we define*

$$rec_X : H_1^S(X, \mathbb{Z}/m\mathbb{Z}) \longrightarrow \pi_1^{t,ab}(X)/m$$

as the homomorphism induced by the pairing of Proposition 2.12 for $A = \mathbb{Z}/m\mathbb{Z}$ combined with the isomorphism $H_t^1(X, \mathbb{Z}/m\mathbb{Z})^* \cong \pi_1^{t,ab}(X)/m$.

The statement of the next lemma immediately follows from the definition of rec.

LEMMA 2.14. *Let $f : X' \to X$ be a morphism of separated schemes of finite type over k. Then the induced diagram*

$$\begin{array}{ccc}
H_1^S(X', \mathbb{Z}/m\mathbb{Z}) & \xrightarrow{rec_{X'}} & \pi_1^{t,ab}(X')/m \\
\downarrow f_* & & \downarrow f_* \\
H_1^S(X, \mathbb{Z}/m\mathbb{Z}) & \xrightarrow{rec_X} & \pi_1^{t,ab}(X)/m
\end{array}$$

commutes.

3 RIGID ČECH COMPLEXES

We consider étale sheaves F on the category Sch/k of separated schemes of finite type over a field k. By a result of M. Artin, Čech cohomology $\check{H}^\bullet(X, F)$ and sheaf cohomology $H^\bullet_{\text{et}}(X, F)$ coincide in degree ≤ 1 and in arbitrary degree if X is quasi-projective (cf. [Mi], III Thm. 2.17). Comparing the Čech complex for a covering $\mathcal{U}$ and that for a finer covering $\mathcal{V}$, the refinement homomorphism

$$\check{C}^\bullet(\mathcal{U}, F) \longrightarrow \check{C}^\bullet(\mathcal{V}, F)$$

is canonical only up to chain homotopy and hence only the induced map $\check{H}^\bullet(\mathcal{U},F) \to \check{H}^\bullet(\mathcal{V},F)$ is well-defined. We can remedy this problem in the spirit of Friedlander [Fr], chap.4, by using rigid coverings:
We fix an algebraic closure $\bar{k}/k$. A *rigid étale covering* $\mathcal{U}$ of X is a family of pointed separated étale morphisms

$$(U_x,u_x)\longrightarrow (X,x),\quad x\in X(\bar{k}),$$

with U_x connected and $u_x \in U_x(\bar{k})$ mapping to x. For an étale sheaf F the rigid Čech complex is defined by

$$\check{C}^\bullet(\mathcal{U},F):\quad \check{C}^n(\mathcal{U},F)=\prod_{(x_0,\ldots,x_n)\in X(\bar{k})^{n+1}}\Gamma(U_{x_0}\times_X\cdots\times_X U_{x_n},F)$$

with the usual differentials. It is clear what it means for a rigid covering $\mathcal{V}$ to be a refinement of $\mathcal{U}$. Because the marked points map to each other, there is exactly one refinement morphism, hence we obtain a canonical refinement morphism on the level of complexes

$$\check{C}^\bullet(\mathcal{U},F)\to\check{C}^\bullet(\mathcal{V},F).$$

The set of rigid coverings is cofiltered (form the fibre product for each $x \in X(\bar{k})$ and restrict to the connected components of the marked points). Therefore we can define the rigid Čech complex of X with values in F as the filtered direct limit

$$\check{C}^\bullet(X,F):=\varinjlim_{\mathcal{U}}\check{C}^\bullet(\mathcal{U},F),$$

where $\mathcal{U}$ runs through all rigid coverings of X. Note that the rigid Čech complex depends on the structure morphism $X \to k$ and not merely on the scheme X.
Forgetting the marking, we can view a rigid covering as a usual covering. Every covering can be refined by a covering which arises by forgetting the marking of a rigid covering. Hence the cohomology of the rigid Čech complex coincides with the usual Čech cohomology of X with values in F.

For a morphism $f : Y \to X$ and a rigid Čech covering $\mathcal{U}/X$, we obtain a rigid Čech covering $f^*\mathcal{U}/Y$ by taking base extension to Y and restricting to the connected components of the marked points, and in the limit we obtain a homomorphism

$$f^*:\check{C}^\bullet(X,F)\longrightarrow\check{C}^\bullet(Y,F).$$

LEMMA 3.1. *If $\pi : Y \to X$ is quasi-finite, then the rigid coverings of the form $\pi^*\mathcal{U}$ are cofinal among the rigid coverings of Y.*

Proof. This is an immediate consequence of the fact that a quasi-finite and separated scheme Y over the spectrum X of a henselian ring is of the form $Y = Y_0 \sqcup Y_1 \sqcup \ldots \sqcup Y_r$ with $Y_0 \to X$ not surjective and $Y_i \to X$ finite surjective with Y_i the spectrum of a henselian ring, $i = 1,\ldots,r$, cf. [Mi], I, Thm. 4.2. □

LEMMA 3.2. *If F is qfh-sheaf on Sch/k, then for any $n \geq 0$ the presheaf $\underline{\check{C}}^n(-,F)$ given by*

$$X \longmapsto \check{C}^n(X,F)$$

is a qfh-sheaf. The obvious sequence

$$0 \to F \to \underline{\check{C}}^0(-,F) \to \underline{\check{C}}^1(-,F) \to \underline{\check{C}}^2(-,F) \to \cdots$$

is exact as a sequence of étale (and hence also of qfh) sheaves.

Proof. We show that each $\underline{\check{C}}^n(-,F)$ is a qfh-sheaf. For this, let $\pi : Y \to X$ be a qfh-covering, i.e., a quasi-finite universal topological epimorphism. We denote the projection by $\Pi : Y \times_X Y \to X$. By Lemma 3.1, we have to show that the sequence

$$\varinjlim_{\mathcal{U}} \check{C}^n(\mathcal{U},F) \to \varinjlim_{\mathcal{U}} \check{C}^n(\pi^*\mathcal{U},F) \rightrightarrows \varinjlim_{\mathcal{U}} \check{C}^n(\Pi^*\mathcal{U},F)$$

is an equalizer, where $\mathcal{U}$ runs through the rigid coverings of X. Since filtered colimits commute with finite limits, it suffices to show the exactness for a single, sufficiently small $\mathcal{U}$. This, however, follows from the assumption that F is a qfh-sheaf.
Finally, the exactness of $0 \to F \to \underline{\check{C}}^0(-,F) \to \underline{\check{C}}^1(-,F) \to \cdots$ as a sequence of étale sheaves follows by considering stalks. □

Being qfh-sheaves, the sheaves F and $\underline{\check{C}}^n(-,F)$ admit transfer maps, see [SV1], §5. For later use, we make the relation between the transfers of F and of $\underline{\check{C}}^n(-,F)$ explicit: Let Z be integral, X regular and $\pi : Z \to X$ finite and surjective. Let F be a qfh-sheaf on Sch/k. For $x \in X(\bar{k})$ we have

$$X_x^{sh} \times_X Z = \coprod_{z \in \pi^{-1}(x)} Z_z^{sh},$$

where $\pi^{-1}(x)$ denotes the set of morphisms $z : \mathit{Spec}(\bar{k}) \to Z$ with $\pi \circ z = x$. For sufficiently small étale $(U_x,u_x) \to (X,x)$, the set of connected components of $U_x \times_X Z$ is in 1-1-correspondence with the set $\pi^{-1}(x)$, and to each family of étale morphisms

$$(V_z,v_z) \longrightarrow (Z,z), \quad z \in \pi^{-1}(x),$$

there is (after possibly making U_x smaller) a unique morphism

$$U_x \times_X Z \longrightarrow \coprod_{z \in \pi^{-1}(x)} V_z,$$

over Z, which sends the connected component associated with z of $U_x \times_X Z$ to V_z, and the point (u_x,z) to v_z.

In this way we obtain, for finitely many points $(x_0,\ldots,x_n)$, $n \geq 0$, and for every family

$$(V_{z_i,v_{z_i}}) \longrightarrow (Z,z_i), \quad z_i \in \pi^{-1}(x_i),$$

and sufficiently small chosen

$$(U_{x_i}, u_{x_i}) \longrightarrow (X, x_i), \quad i = 0, \dots, n,$$

a homomorphism

$$\prod_{\substack{(z_0,\dots,z_n)\\ z_i\in\pi^{-1}(x_i)}} \Gamma\big(V_{z_0}\times_Z\cdots\times_Z V_{z_n}, F\big) \longrightarrow \Gamma\big(U_{x_0}\times_X\cdots\times_X U_{x_n}\times_X Z, F\big).$$

Since F is a qfh-sheaf, we can compose this with the transfer map associated with the finite morphism

$$U_{x_0}\times_X\cdots\times_X U_{x_n}\times_X Z \to U_{x_0}\times_X\cdots\times_X U_{x_n}.$$

Forming for fixed n the product over all $(x_0,\dots,x_n)\in X(\bar{k})^{n+1}$ and passing to the limit over all rigid coverings, we obtain the transfer homomorphism

$$\pi_* : \check{C}^\bullet(Z,F) \longrightarrow \check{C}^\bullet(X,F).$$

Passing to cohomology, we obtain the usual transfer on étale cohomology in degree 0 and 1, and in any degree if the schemes are quasi-projective.

Next we give the pairing

$$\langle\cdot,\cdot\rangle : H_1^S(X,\mathbb{Z}/m\mathbb{Z})\times H_t^1(X,A) \longrightarrow A.$$

constructed in Proposition 2.12 for k algebraically closed and an abelian group A with $mA=0$ the following interpretation in terms of the rigid Čech complex:

Let $a\in H_1^S(X,\mathbb{Z}/m\mathbb{Z})$ and $b\in H_t^1(X,A)$ be given, and let $\alpha\in\mathrm{Cor}_k(\Delta^1,X)$ and $\beta\in\ker(\check{C}^1(X,A)\xrightarrow{d}\check{C}^2(X,A))$ be representing elements. Note that $(0^*-1^*)(\alpha)\in m\,\mathrm{Cor}(\Delta^0,X)$ by assumption. Consider the diagram

$$\begin{array}{ccccccc}
 & & \check{C}^0(X,A) & \xrightarrow{d} & \check{C}^1(X,A) & \xrightarrow{d} & \check{C}^2(X,A)\\
 & & \downarrow{\scriptstyle\alpha^*} & & \downarrow{\scriptstyle\alpha^*} & & \downarrow{\scriptstyle\alpha^*}\\
 & & \check{C}^0(\Delta^1,A) & \xrightarrow{d} & \check{C}^1(\Delta^1,A) & \xrightarrow{d} & \check{C}^2(\Delta^1,A)\\
 & & \downarrow{\scriptstyle 0^*-1^*} & & \downarrow{\scriptstyle 0^*-1^*} & & \downarrow{\scriptstyle 0^*-1^*}\\
A & \hookrightarrow & \check{C}^0(\Delta^0,A) & \xrightarrow{d} & \check{C}^1(\Delta^0,A) & \xrightarrow{d} & \check{C}^2(\Delta^0,A)
\end{array}$$

Since β represents a tame torsor $\mathcal{T}$ on X, $\alpha^*(\beta)$ represents the torsor $\alpha^*(\mathcal{T})$, which is tame by Lemma 2.9. By Corollary 2.11, there exists $\gamma\in\check{C}^0(\Delta^1,A)$ with $d\gamma=\alpha^*(\beta)$. Since

$$d(0^*-1^*)(\gamma) = (0^*-1^*)\alpha^*(\beta) = 0,$$

we conclude that $(0^* - 1^*)(\gamma)$ lies in

$$A = H^0(\Delta^0, A) = \ker(\check{C}^0(\Delta^0, A) \xrightarrow{d} \check{C}^1(\Delta^0, A)).$$

It is easy to verify that the assignment

$$\langle \cdot, \cdot \rangle : (a, b) \longmapsto (0^* - 1^*)(\gamma) \in A$$

does not depend on the choices made. By the explicit geometric relation between Čech 1-cocycles and torsors, and since our construction of finite push-forwards of torsors is compatible with the construction of transfers for qfh-sheaves given in [SV1], §5, we see that the pairing constructed above coincides with the one constructed in Proposition 2.12.

Finally, let

$$A \hookrightarrow I^0 \to I^1 \to I^2 \tag{5}$$

be a (partial) injective resolution of the constant sheaf A in the category of $\mathbb{Z}/m\mathbb{Z}$-module sheaves on $(\mathrm{Sch}/k)_{\mathrm{qfh}}$. Let $\phi : (\mathrm{Sch}/k)_{\mathrm{qfh}} \to (\mathrm{Sch}/k)_{\mathrm{et}}$ denote the natural map of sites. Since ϕ^* is exact, ϕ_* sends injective sheaves to injective sheaves. By [SV1], Thm. 10.2, we have $R^0\phi_*(A) = A$ and $R^i\phi_*(A) = 0$ for $i \geq 1$. Hence (5) is also a partial resolution of A by injective, étale sheaves of $\mathbb{Z}/m\mathbb{Z}$-modules. We choose a quasi-isomorphism

$$[0 \to \underline{\check{C}}^0(-, A) \to \underline{\check{C}}^1(-, A) \to \underline{\check{C}}^2(-, A)] \longrightarrow [0 \to I^0 \to I^1 \to I^2]$$

of truncated complexes of qfh-sheaves. Since Čech- and étale cohomology agree in dimension ≤ 1, the induced map on global sections is a quasi-isomorphism of truncated complexes of abelian groups. Hence the pairing of Proposition 2.12 can also be obtained by the same procedure as above but using the diagram

$$\begin{tikzcd}
 & I^0(X) \arrow[r, "d"] \arrow[d, "\alpha^*"] & I^1(X) \arrow[r, "d"] \arrow[d, "\alpha^*"] & I^2(X) \arrow[d, "\alpha^*"] \\
 & I^0(\Delta^1) \arrow[r, "d"] \arrow[d, "0^* - 1^*"] & I^1(\Delta^1) \arrow[r, "d"] \arrow[d, "0^* - 1^*"] & I^2(\Delta^1) \arrow[d, "0^* - 1^*"] \\
A \arrow[r, hook] & I^0(\Delta^0) \arrow[r, "d"] & I^1(\Delta^0) \arrow[r, "d"] & I^2(\Delta^0).
\end{tikzcd}$$

By [SV1], Theorem 10.7, the same argument applies with a partial injective resolution of the constant sheaf A in the category of $\mathbb{Z}/m\mathbb{Z}$-module sheaves on $(\mathrm{Sch}/k)_h$.

4 THE CASE OF SMOOTH CURVES

In this section we prove Theorem 1.1 in the case that $X = C$ is a smooth curve.

Let k be an algebraically closed field of characteristic $p \geq 0$, and let C be a smooth, but not necessarily projective, curve over k. Let the semi-abelian variety $\mathcal{A}$ be the generalized Jacobian of C with respect to the modulus given by the sum of the points on the boundary of the regular compactification $\bar{C}$ of C (cf. [Se], Ch. 5). The group $\mathcal{A}(k)$ is the subgroup of degree zero elements of the relative Picard group $\mathrm{Pic}(\bar{C}, \bar{C} \smallsetminus C)$. By [SV1], Thm. 3.1 (see [Li], for the case $C = \bar{C}$), there is an isomorphism

$$H_0^S(C,\mathbb{Z})^0 := \ker(H_0^S(C,\mathbb{Z}) \xrightarrow{\deg} \mathbb{Z}) \cong \mathcal{A}(k),$$

in particular, $\mathcal{A}(k)$ is a quotient of the group of zero cycles of degree zero on C. From the coefficient sequence together with the divisibility of $H_1^S(C,\mathbb{Z})$ (which is isomorphic to $k^\times$ if C is proper and zero otherwise), we obtain an isomorphism

$$H_1^S(C,\mathbb{Z}/m\mathbb{Z}) \underset{\sim}{\xrightarrow{\delta}} {}_mH_0^S(C,\mathbb{Z}) \cong {}_m\mathcal{A}(k). \tag{6}$$

After fixing a closed point P_0 of C, the morphism $C \to \mathcal{A}$, $P \mapsto P - P_0$, is universal for morphisms of C to semi-abelian varieties, i.e., $\mathcal{A}$ is the generalized Albanese variety of C ([Se], V, Th. 2).

Consider the m-multiplication map $\mathcal{A} \stackrel{m}{\to} \mathcal{A}$. Its maximal étale subcovering $\widetilde{\mathcal{A}} \to \mathcal{A}$ is the quotient of $\mathcal{A}$ by the connected component of the finite group scheme ${}_m\mathcal{A}$ (if $(p,m) = 1$, the connected component is trivial). The projection $\mathcal{A} \to \widetilde{\mathcal{A}}$ induces an isomorphism $\mathcal{A}(k) \stackrel{\sim}{\to} \widetilde{\mathcal{A}}(k)$ on rational points, and we identify $\mathcal{A}(k)$ and $\widetilde{\mathcal{A}}(k)$ via this isomorphism. With respect to this identification, the projection $\widetilde{A}(k) \to \mathcal{A}(k)$ is the m-multiplication map on $\mathcal{A}(k)$.

By [Se], Ch. IV, $\widetilde{C} := C \times_{\mathcal{A}} \widetilde{\mathcal{A}}$ is the maximal abelian tame étale covering of C with Galois group annihilated by m. Because $\mathrm{Aut}_{\mathcal{A}}(\widetilde{\mathcal{A}}) \cong {}_m\mathcal{A}(k)$, we obtain an isomorphism

$$\mathrm{Hom}({}_m\mathcal{A}(k), A) \underset{\sim}{\xrightarrow{\tau}} H_t^1(C,A) \tag{7}$$

for any finite abelian group A with $mA = 0$.

THEOREM 4.1. *For any finite abelian group A with $mA = 0$, the diagram*

$$\begin{array}{ccccc}
H_1^S(C,\mathbb{Z}/m\mathbb{Z}) & \times & H_t^1(C,A) & \xrightarrow{\langle\ ,\ \rangle} & A \\
\delta \downarrow \wr & & \wr \uparrow \tau & & \| \\
{}_m\mathcal{A}(k) & \times & \mathrm{Hom}({}_m\mathcal{A}(k), A) & \xrightarrow{\text{eval}} & A
\end{array}$$

where $\langle\ ,\ \rangle$ is the pairing from Proposition 2.12 and $eval$ is the evaluation map, commutes. In particular, the upper pairing is perfect and the induced homomorphism $H_1^S(C,\mathbb{Z}/m\mathbb{Z}) \to \pi_1^{t,ab}(C)/m$ is an isomorphism.

Proof. We have to show that $\phi(\delta(\zeta)) = \langle\zeta, \tau(\phi)\rangle$ for any $\zeta \in H_1^S(C, \mathbb{Z}/m\mathbb{Z})$ and any $\phi \in \operatorname{Hom}({}_m\mathcal{A}(k), A)$. By functoriality, it suffices to consider the universal case $A = {}_m\mathcal{A}(k)$, $\phi = \mathrm{id}$. In this case $\tau(\mathrm{id})$ is the torsor $\widetilde{\pi} : \widetilde{C} \to C$.

Let C' be the regular compactification of C. By [SV1], Thm. 3.1, $\delta(\zeta) \in {}_mH_0^S(C, \mathbb{Z}) = {}_m\mathcal{A}(k)$ is the class $[z]$ of some $z \in Z_0(C)$ (the group of zero-cycles on C) such that

$$mz = \gamma^*(0) - \gamma^*(1)$$

for some finite morphism $\gamma : C' \to \mathbb{P}^1$ with $C' \smallsetminus C \subset \gamma^{-1}(\infty)$. The diagram

$$\begin{tikzcd} C' \smallsetminus \gamma^{-1}(\infty) \arrow[r, hook] \arrow[d, "\gamma|_{C' \smallsetminus \gamma^{-1}(\infty)}"] & C \arrow[r, hook] & C' \arrow[d, "\gamma"] \\ \Delta^1 \arrow[rr, hook] & & \mathbb{P}^1 \end{tikzcd}$$

shows that γ induces a finite correspondence, say g, from Δ^1 to C. The class of g in $H_1^S(C, \mathbb{Z}/m\mathbb{Z})$ is a pre-image of $\delta(\zeta)$ under $H_1^S(C, \mathbb{Z}/m\mathbb{Z}) \xrightarrow{\sim} {}_mH_0^S(C, \mathbb{Z})$, i.e., ζ is represented by g. It therefore suffices to show that

$$[z] = \langle g, \widetilde{C}\rangle.$$

Let d be the degree of γ and $\gamma^*(0) = \sum_{i=1}^d P_i$, $\gamma^*(1) = \sum_{i=1}^d Q_i$. Each point in $\gamma^*(0)$ and $\gamma^*(1)$ occurs with multiplicity divisible by m, in particular $d = mr$ for some integer r. After reindexing, we may assume that $P_i = P_j$ and $Q_i = Q_j$ for $i \equiv j \bmod r$, hence

$$z = \sum_{i=1}^r P_i - \sum_{i=1}^r Q_i.$$

On the level of closed points, $\widetilde{C} = C \times_{\mathcal{A}} \widetilde{\mathcal{A}}$ can be identified with the set of $a \in \widetilde{\mathcal{A}}(k) = \mathcal{A}(k)$ such that $ma = P - P_0$ for some point $P \in C$ (a projects to P in C, i.e., $\widetilde{\pi}(a) = P$). The ${}_m\mathcal{A}(k)$-principal homogeneous space $0^*g^*\widetilde{C}$ can be identified with the quotient of the set

$$\prod_{i=1}^d \widetilde{\pi}^{-1}(P_i)$$

by the action of ${}_m\mathcal{A}(k)^{d-1}$ given by

$$(\beta_1, \ldots, \beta_{d-1})(a_1, \ldots, a_d) = (a_1 + \beta_1, a_2 - \beta_1 + \beta_2, \ldots, a_d - \beta_{d-1}).$$

We fix points $a_1, \ldots, a_d \in \widetilde{C}$ over $P_1, \ldots, P_d$ subject to the condition $a_i = a_j$ for $P_i = P_j$. Then $0^*g^*\widetilde{C}$ is identified with the quotient of the set

$$(a_1 + {}_m\mathcal{A}(k)) \times \cdots \times (a_d + {}_m\mathcal{A}(k))$$

by the action of ${}_m\mathcal{A}(k)^{d-1}$. Since each a_i occurs with multiplicity divisible by m, the trivialization $0^*g^*(\widetilde{C}) \xrightarrow{\sim} {}_m\mathcal{A}(k)$ given by

$$\overline{(a_1 + \alpha_1, \ldots, a_d + \alpha_d)} \longmapsto \alpha_1 + \cdots + \alpha_d \in {}_m\mathcal{A}(k)$$

does not depend on the choice of the a_i. We do the same with $1^*g^*(\widetilde{C})$ by choosing $b_i \in \widetilde{C}$ over Q_i. Then we see that the tautological identification $\Phi_{taut} : 0^*g^*(\widetilde{C}) \xrightarrow{\sim} 1^*g^*(\widetilde{C})$ is given by

$$\overline{(a_1+\alpha_1,\ldots,a_d+\alpha_d)} \longmapsto \overline{(b_1+\alpha_1,\ldots,b_d+\alpha_d)}.$$

Now consider the morphism

$$\Sigma : \mathrm{Sym}^d(C) \longrightarrow \mathcal{A},\ (x_1,\ldots,x_d) \longmapsto [\sum(x_i - P_0)].$$

Associated with the ${}_m\mathcal{A}(k)$-torsor $\widetilde{C}$ over C, we have the ${}_m\mathcal{A}(k)$-torsor $\mathrm{Sym}^d_{{}_m\mathcal{A}(k)}(\widetilde{C})$ over $\mathrm{Sym}^d(C)$ (cf. the paragraph preceding Lemma 2.3). The commutative diagram

$$\begin{array}{ccc} \mathrm{Sym}^d_{{}_m\mathcal{A}(k)}(\widetilde{C}) & \xrightarrow{\ \Sigma\ } & \widetilde{\mathcal{A}} \\ \downarrow & & \downarrow \\ \mathrm{Sym}^d(C) & \xrightarrow{\ \Sigma\ } & \mathcal{A} \end{array}$$

induces a map (hence an isomorphism) of ${}_m\mathcal{A}(k)$-torsors $\mathrm{Sym}^d_{{}_m\mathcal{A}(k)}(\widetilde{C}) \xrightarrow{\sim} \widetilde{\mathcal{A}} \times_{\mathcal{A}} \mathrm{Sym}^d(C)$. Consider the morphism $S_g : \Delta^1_k \to \mathrm{Sym}^d(C)$ associated with the finite correspondence g. Since the generalized Jacobian of $\Delta^1_k \cong \mathbb{A}^1_k$ is $Spec(k)$, the composite

$$\Delta^1_k \xrightarrow{S_g} \mathrm{Sym}^d(C) \xrightarrow{\Sigma} \mathcal{A}$$

is constant with value $a := [\sum_{i=1}^d (P_i - P_0)] = [\sum_{i=1}^d (Q_i - P_0)] \in \mathcal{A}(k)$. By Lemma 2.3, we obtain an isomorphism

$$g^*(\widetilde{C}) = S_g^*(\mathrm{Sym}^d(\widetilde{C})) = \Sigma^* S_g^* \widetilde{\mathcal{A}} = \Delta^1_k \times \widetilde{\pi}^{-1}(a)$$

(giving a trivialization after choosing a point in $\widetilde{\pi}^{-1}(a)$). On the fibre over 0 it is given by

$$\overline{(a_1+\alpha_1,\ldots,a_d+\alpha_d)} \longmapsto \sum_{i=1}^d (a_i+\alpha_i) \in \widetilde{\pi}^{-1}(a) \subset \widetilde{\mathcal{A}}$$

and similarly on the fibre over 1. We conclude that $\Phi_{par} \circ \Phi_{taut}^{-1}$ is translation by

$$\sum_{i=1}^d (a_i - b_i) = \sum_{i=1}^r m(a_i - b_i) = \sum_{i=1}^r [P_i - Q_i] = [z].$$

This concludes the proof. □

5 THE BLOW-UP SEQUENCES

All schemes in this section are separated schemes of finite type over the spectrum of a perfect field k. A curve on a scheme X is a closed one-dimensional subscheme. The normalization of a curve C is denoted by $\widetilde{C}$.

Now let

$$\begin{tikzcd} Z' \arrow[r] \arrow[d] & X' \arrow[d, "\pi"] \\ Z \arrow[r, "i"] & X \end{tikzcd}$$

be an abstract blow-up square, i.e., a cartesian diagram of schemes such that $\pi : X' \to X$ is proper, $i : Z \to X$ is a closed embedding and π induces an isomorphism $(X' \smallsetminus Z')_{\text{red}} \xrightarrow{\sim} (X \smallsetminus Z)_{\text{red}}$.

PROPOSITION 5.1. *Given an abstract blow-up square and an abelian group A, assume that π is finite or A is torsion. Then there is a natural exact sequence*

$$0 \to H^0_{\text{et}}(X, A) \to H^0_{\text{et}}(X', A) \oplus H^0_{\text{et}}(Z, A) \to H^0_{\text{et}}(Z', A)$$
$$\xrightarrow{\delta} H^1_t(X, A) \to H^1_t(X', A) \oplus H^1_t(Z, A) \to H^1_t(Z', A).$$

Proof. We call an abstract blow-up square trivial, if i is surjective (i.e., s_{red} is an isomorphism) or if $\pi_{\text{red}} : X'_{\text{red}} \to X_{\text{red}}$ has a section. Every abstract blow-up square with X a connected regular curve is trivial.
Now let an arbitrary abstract blow-up square be given. If A is torsion, the proper base change theorem implies (cf. [Ge2], 3.2 and 3.6) that we have a long exact sequence

$$\cdots \to H^i_{\text{et}}(X, A) \to H^i_{\text{et}}(X', A) \oplus H^i_{\text{et}}(Z, A) \to H^i_{\text{et}}(Z', A) \to H^{i+1}_{\text{et}}(X, A) \to \cdots$$

If π is finite, the same is true for arbitrary A since π_* is exact. If the blow-up square is trivial, this long exact sequence splits into short exact sequences $0 \to H^i_{\text{et}}(X, A) \to H^i_{\text{et}}(X', A) \oplus H^i_{\text{et}}(Z, A) \to H^i_{\text{et}}(Z', A) \to 0$ for all i.

Next we show the exact sequence of the proposition. We omit the coefficients A and put $H^0_t(X) = H^0_{\text{et}}(X)$. We first show, that the image of the boundary map $\delta : H^0_{\text{et}}(Z') \to H^1_{\text{et}}(X)$ has image in $H^1_t(X)$, thus showing the existence of $H^0_t(Z') \to H^1_t(X)$ and, at the same time, the exactness of the sequence at $H^1_t(X)$. Let $\widetilde{C} \to X$ be the normalization of a curve in X. The base change

$$\begin{tikzcd} Z'_{\widetilde{C}} \arrow[r] \arrow[d] & X'_{\widetilde{C}} \arrow[d, "\widetilde{\pi}"] \\ Z_{\widetilde{C}} \arrow[r, "\widetilde{i}"] & \widetilde{C} \end{tikzcd}$$

of our abstract blow-up square to $\widetilde{C}$ is a trivial abstract blow-up square. Therefore, for any $\alpha \in H^0_{\text{et}}(Z')$, the pull-back of α to $H^0_{\text{et}}(Z'_{\widetilde{C}})$ lies in the image of $H^0_{\text{et}}(X'_{\widetilde{C}}) \oplus H^0_{\text{et}}(Z_{\widetilde{C}}) \to H^0_{\text{et}}(Z'_{\widetilde{C}})$ and has therefore trivial image under $\delta : H^0_{\text{et}}(Z'_{\widetilde{C}}) \to H^1_{\text{et}}(\widetilde{C})$.

Therefore, $\delta(\alpha) \in H^1_{\mathrm{et}}(X)$ has trivial image in $H^1_{\mathrm{et}}(\widetilde{C})$ for every curve $C \subset X$, in particular, it lies in $H^1_t(X)$.
It remains to show exactness at $H^1_t(X') \oplus H^1_t(Z)$. Let α be in this group with trivial image in $H^1_t(Z')$. Then there exists $\beta \in H^1_{\mathrm{et}}(X)$ mapping to α and it remains to show that β lies in the subgroup $H^1_t(X)$. But this is clear, because for every curve $C \subset X$ we have $H^1_t(\widetilde{C}) = \ker(H^1_t(X'_{\widetilde{C}}) \oplus H^1_t(Z_{\widetilde{C}}) \to H^1_t(Z'_{\widetilde{C}}))$. □

PROPOSITION 5.2. *Given an abstract blow-up square*

$$\begin{array}{ccc} Z' & \longrightarrow & X' \\ \downarrow & & \downarrow{\scriptstyle \pi} \\ Z & \xrightarrow{i} & X \end{array}$$

and an abelian group A, there is a natural exact sequence of Suslin homology groups

$$H_1^S(Z',A) \to H_1^S(X',A) \oplus H_1^S(Z,A) \to H_1^S(X,A)$$
$$\xrightarrow{\delta} H_0^S(Z',A) \to H_0^S(X',A) \oplus H_0^S(Z,A) \to H_0^S(X,A) \to 0.$$

Proof. Consider the exact sequences

$$C_\bullet(Z',A) \hookrightarrow C_\bullet(X',A) \oplus C_\bullet(Z,A) \to C_\bullet(X,A) \twoheadrightarrow K_\bullet^A$$

and

$$C_\bullet(Z') \hookrightarrow C_\bullet(X') \oplus C_\bullet(Z) \to C_\bullet(X) \twoheadrightarrow K_\bullet,$$

where $K_\bullet^A$ and $K_\bullet$ are defined to make the sequences exact. Since the complexes $C_\bullet(-)$ consist of free abelian groups, in order the show the statement of the proposition, it suffices to show that $H_i(K_\bullet) = 0$ for $i \leq 2$. Let Sm/k be the full subcategory of Sch/k consisting of smooth schemes. For $Y \in \mathrm{Sch}/k$ we consider the presheaf $c(Y)$ on Sm/k given by $c(Y)(U) = \mathrm{Cor}(U,Y)$. Then, by [SV2], Thm. 5.2, 4.7 and its proof, the sequence

$$0 \to c(Z') \to c(X') \oplus c(Z) \overset{(\pi_*,i_*)}{\to} c(X)$$

is exact and $F := \mathit{coker}(\pi_*, i_*)$ has the property that, for any $U \in \mathrm{Sm}/k$ of dimension ≤ 2 and any $x \in F(U)$, there exists a proper birational morphism $\phi : V \to U$ with V smooth such that $\phi^*(x) = 0$. Let $F_\bullet$ be the complex of presheaves given by $F_n(U) = F(U \times \Delta^n)$ with the obvious differentials and let $(F_\bullet)_{\mathrm{Nis}}$ be the associated complex of sheaves on $(\mathrm{Sm}/k)_{\mathrm{Nis}}$. Then by [SS], Thm. 2.4, the Nisnevich sheaves

$$\mathcal{H}_i((F_\bullet)_{\mathrm{Nis}})$$

vanish for $i \leq 2$. Evaluating at $U = \mathit{Spec}(k)$ yields the result. □

Now assume that k is algebraically closed. Let

$$rec_{1,X} : H_1^S(X, \mathbb{Z}/m\mathbb{Z}) \to H_t^1(X, \mathbb{Z}/m\mathbb{Z})^*$$

be the reciprocity map constructed in Section 2 and let

$$rec_{0,X} : H_0^S(X, \mathbb{Z}/m\mathbb{Z}) \to H_{\mathrm{et}}^0(X, \mathbb{Z}/m\mathbb{Z})^*$$

be the homomorphism induced by the pairing

$$\langle \cdot, \cdot \rangle : \; H_0^S(X, \mathbb{Z}/m\mathbb{Z}) \times H_{\mathrm{et}}^0(X, \mathbb{Z}/m\mathbb{Z}) \longrightarrow \mathbb{Z}/m\mathbb{Z}$$

defined as follows: Given $a \in H_0^S(X, \mathbb{Z}/m\mathbb{Z})$ and $b \in H_{\mathrm{et}}^0(X, \mathbb{Z}/m\mathbb{Z})$, we represent a by a correspondence $\alpha \in \mathrm{Cor}(\Delta^0, X)$ and put $\langle a, b\rangle = \alpha^*(b) \in H_{\mathrm{et}}^0(\Delta^0, \mathbb{Z}/m\mathbb{Z}) \cong \mathbb{Z}/m\mathbb{Z}$. This is well-defined since the homomorphisms $0^*, 1^* : H_{\mathrm{et}}^0(\Delta^1, \mathbb{Z}/m\mathbb{Z}) \to H_{\mathrm{et}}^0(\Delta^0, \mathbb{Z}/m\mathbb{Z})$ agree.

LEMMA 5.3. *For any m, $rec_{0,X}$ is an isomorphism.*

Proof. For connected X, we have the commutative diagram

$$\begin{array}{ccc} H_0^S(X, \mathbb{Z}/m\mathbb{Z}) & \xrightarrow{rec_{0,X}} & H_{\mathrm{et}}^0(X, \mathbb{Z}/m\mathbb{Z})^* \\ \big\downarrow \mathrm{deg} & & \big\downarrow \wr \\ H_0^S(k, \mathbb{Z}/m\mathbb{Z}) & \xrightarrow[\sim]{rec_{0,k}} & H_{\mathrm{et}}^0(k, \mathbb{Z}/m\mathbb{Z})^*. \end{array}$$

Hence, for connected X, it suffices by functoriality to consider the mod m degree map. In particular, $rec_{0,X}$ is surjective for arbitrary X and is an isomorphism if $\dim X = 0$. If X is a smooth connected curve, then $H_0^S(X, \mathbb{Z}) = \mathrm{Pic}(\bar{X}, \bar{X} \smallsetminus X)$, where $\bar{X}$ is the smooth compactification of X (cf. [SV1], Thm. 3.1). The subgroup $\mathrm{Pic}^0(\bar{X}, \bar{X} \smallsetminus X)$ of degree zero elements is the group of k-rational points of the Albanese of X, and hence divisible. Therefore, $rec_{0,X}$ is an isomorphism for connected, and hence for all smooth curves. Considering the normalization morphism of an arbitrary scheme of dimension 1 and the exact sequences of Propositions 5.1 and 5.2, the five-lemma shows that $rec_{0,X}$ is a isomorphism for $\dim X \le 1$.
It remains to show that $rec_{0,X}$ is injective for arbitrary X. We may assume X to be connected. Let $a \in \ker(rec_{0,X})$ and let $\alpha \in Z_0(X)$ be a representing 0-cycle. Since $supp(\alpha)$ is finite, we can find a connected 1-dimensional closed subscheme $Z \subset X$ containing $supp(\alpha)$ (use, e.g., [Mu], II §6 Lemma). Since $rec_{0,Z}$ is injective and a is in the image of $H_0^S(Z, \mathbb{Z}/m\mathbb{Z}) \to H_0^S(X, \mathbb{Z}/m\mathbb{Z})$, we conclude that $a = 0$. □

COROLLARY 5.4. *Let k be an algebraically closed field and let $X \in \mathrm{Sch}/k$ be connected. Then the kernel of the degree map*

$$\deg : H_0^S(X, \mathbb{Z}) \longrightarrow H_0^S(k, \mathbb{Z}) \cong \mathbb{Z}$$

is divisible.

PROPOSITION 5.5. *Let k be algebraically closed and let*

$$\begin{tikzcd} Z' \arrow[r, "i'"] \arrow[d, "\pi'"] & X' \arrow[d, "\pi"] \\ Z \arrow[r, "i"] & X \end{tikzcd}$$

be an abstract blow-up square. Then for any integer $m \geq 1$ the diagram

$$\begin{tikzcd} H_1^S(X,\mathbb{Z}/m\mathbb{Z}) \arrow[r, "\delta"] \arrow[d, "rec_{1,X}"] & H_0^S(Z',\mathbb{Z}/m\mathbb{Z}) \arrow[d, "rec_{0,X}"] \\ H_t^1(X,\mathbb{Z}/m\mathbb{Z})^* \arrow[r, "\delta^*"] & H_{\mathrm{et}}^0(Z',\mathbb{Z}/m\mathbb{Z})^*, \end{tikzcd}$$

commutes. Here δ is the boundary map of Proposition 5.2 and δ^* is the dual of the boundary map of Proposition 5.1.

Proof. We have to show that the diagram

$$\begin{tikzcd} H_1^S(X,\mathbb{Z}/m\mathbb{Z}) \arrow[d, "\delta"'] & \times & H_t^1(X,\mathbb{Z}/m\mathbb{Z}) \arrow[r, "\langle\,,\,\rangle"] & \mathbb{Z}/m\mathbb{Z} \arrow[d, equal] \\ H_0^S(Z',\mathbb{Z}/m\mathbb{Z}) & \times & H_{\mathrm{et}}^0(Z',\mathbb{Z}/m\mathbb{Z}) \arrow[u, "\delta"] \arrow[r, "\langle\,,\,\rangle"] & \mathbb{Z}/m\mathbb{Z} \end{tikzcd}$$

commutes. Given $a \in H_1^S(X,\mathbb{Z}/m\mathbb{Z})$ and $b \in H_{\mathrm{et}}^0(Z',\mathbb{Z}/m\mathbb{Z})$, we choose a representing correspondence $\alpha \in C_1(X,\mathbb{Z}/m\mathbb{Z}) = \mathrm{Cor}(\Delta^1, X) \otimes \mathbb{Z}/m\mathbb{Z}$ in such a way that it has a pre-image $\widehat{\alpha} \in C_1(X',\mathbb{Z}/m\mathbb{Z}) \oplus C_1(Z,\mathbb{Z}/m\mathbb{Z})$ (see the proof of Proposition 5.2). By definition, $\delta a \in H_0^S(Z',\mathbb{Z}/m\mathbb{Z})$ is represented by a correspondence $\gamma \in C_0(Z',\mathbb{Z}/m\mathbb{Z})$ such that the diagram

$$\begin{tikzcd} \Delta^0 \arrow[r, "0-1"] \arrow[d, "\gamma"'] & \Delta^1 \arrow[d, "\widehat{\alpha}"] \\ Z' \arrow[r, "i'-\pi'"] & X' \amalg Z \end{tikzcd}$$

of correspondences commutes modulo m. Next choose an injective resolution $\mathbb{Z}/m\mathbb{Z} \to I^\bullet$ of $\mathbb{Z}/m\mathbb{Z}$ in the category of sheaves of $\mathbb{Z}/m\mathbb{Z}$-modules on $(\mathrm{Sch}/k)_h$ in order to compute the pairings (cf. the end of section 3). Consider the following diagram

$$\begin{tikzcd}
 & & I^0(X') \oplus I^0(Z) \arrow[r, "i'^*-\pi'^*"] \arrow[d, "d"] \arrow[dddll, bend right, "\widehat{\alpha}^*"'] & I^0(Z') \arrow[d, "d"] \\
 & I^1(X) \arrow[r, hook, "(\pi^*, i^*)"] \arrow[d, "\alpha^*"] & I^1(X') \oplus I^1(Z) \arrow[r, "i'^*-\pi'^*"] \arrow[d, "\widehat{\alpha}^*"] & I^1(Z') \\
I^0(\Delta^1) \arrow[r, "d"] \arrow[d, "0^*-1^*"] & I^1(\Delta^1) \arrow[r, equal] \arrow[d, "0^*-1^*"] & I^1(\Delta^1) \arrow[r, "d"] \arrow[d, "0^*-1^*"] & I^2(\Delta^1) \arrow[d, "0^*-1^*"] \\
I^0(\Delta^0) \arrow[r, "d"] & I^1(\Delta^0) \arrow[r, equal] & I^1(\Delta^0) \arrow[r, "d"] & I^2(\Delta^0)
\end{tikzcd}$$

By the argument of [MVW] Lemma 12.7, the sequence

$$0 \to F(X) \to F(X') \oplus F(Z) \to F(Z')$$

is exact for every h-sheaf F. Therefore the second line in the diagram is exact. The proper base change theorem implies (cf. [Ge2], 3.2 and 3.6) that

$$I^\bullet(X) \longrightarrow I^\bullet(X') \oplus I^\bullet(Z) \longrightarrow I^\bullet(Z') \xrightarrow{[1]}$$

is an exact triangle in $D(Ab)$. For the exact sequence of complexes

$$0 \to I^\bullet(X) \to I^\bullet(X') \oplus I^\bullet(Z) \to I^\bullet(Z') \to coker^\bullet \to 0,$$

this implies that the complex $coker^\bullet$ is exact. Therefore, $b \in \ker(I^0(Z') \to I^1(Z'))$ has a pre-image $\widehat{\beta} \in I^0(X') \oplus I^0(Z)$. Then

$$d\widehat{\beta} \in \ker(I^1(X') \oplus I^1(Z) \to I^1(Z')),$$

and there exists a unique $\varepsilon \in I^1(X)$ with $(\pi^*, i^*)(\varepsilon) = d\widehat{\beta}$ representing $\delta b \in H^1_t(X)$. We see that $\widehat{\alpha}^*(d\widehat{\beta}) = \alpha^*(\varepsilon)$. It follows that

$$d(\widehat{\alpha}^*(\widehat{\beta})) = \widehat{\alpha}^*(d\widehat{\beta}) = \alpha^*(\varepsilon) \in \ker(I^1(\Delta^1) \xrightarrow{0^*-1^*} I^1(\Delta^0)).$$

By definition of $\langle\ ,\ \rangle$, we obtain

$$\langle a, \delta(b)\rangle = (0^* - 1^*)\widehat{\alpha}^*\widehat{\beta} \in \ker(I^0(\Delta^0) \to I^1(\Delta^0)) = \mathbb{Z}/m\mathbb{Z}.$$

On the other hand, $\langle \delta a, b\rangle = \gamma^*(b) \in H^0_{\mathrm{et}}(\Delta^0)$ is represented by $\gamma^*\beta \in I^0(\Delta^0)$ and the commutative diagram of correspondences above implies

$$\gamma^*\beta = \gamma^*(i'^* - \pi'^*)(\widehat{\beta}) = (0^* - 1^*)\widehat{\alpha}^*\widehat{\beta}.$$

This finishes the proof. □

PROPOSITION 5.6. *Let X be a normal, generically smooth, connected scheme of finite type over a field k and let $M \subseteq H^1_{\mathrm{et}}(X, \mathbb{Z}/m\mathbb{Z})$ be a finite subgroup. Then there exists a regular curve C over k and a finite morphism $\phi : C \to X$ such that M has trivial intersection with the kernel of $\phi^* : H^1_{\mathrm{et}}(X, \mathbb{Z}/m\mathbb{Z}) \to H^1_{\mathrm{et}}(C, \mathbb{Z}/m\mathbb{Z})$.*

Proof. For any normal scheme Z and dense open subscheme $Z' \subset Z$, the induced map $H^1_{\mathrm{et}}(Z, \mathbb{Z}/m\mathbb{Z}) \to H^1_{\mathrm{et}}(Z', \mathbb{Z}/m\mathbb{Z})$ is injective. Hence we may replace X by an open subscheme and assume that X is smooth. Let $Y \to X$ be the finite abelian étale covering corresponding to the kernel of $\pi_1^{ab}(X) \twoheadrightarrow M^*$. We have to find a regular curve C and a finite morphism $C \to X$ such that $C \times_X Y$ is connected.
Choose a separating transcendence basis $t_1, \ldots, t_d$ of $k(X)$ over k. This yields a rational map $X \to \mathbb{P}^d_k$. Let t be another indeterminate and let X_t (resp. Y_t) be the base change of X (resp. Y) to the rational function field $k(t)$. Consider the composition $\phi : Y_t \to X_t \to \mathbb{P}^d_{k(t)}$. Since $k(t)$ is Hilbertian [FJ], Thm. 12.10, we can find a rational point $P \in \mathbb{P}^d_{k(t)}$ over which ϕ is defined and such that P has exactly one pre-image y_t in Y_t. The image $x_t \in X_t$ of y_t has exactly one pre-image in Y_t. Let x be the image of x_t in X. If $\mathrm{trdeg}_k k(x) = 1$ put $x' = x$, if $\mathrm{trdeg}_k k(x) = 0$ (i.e., x is a closed point in X) choose any $x' \in X$ with $\mathrm{trdeg}_k k(x') = 1$ such that x is a regular point of the closure of x'. In both cases the normalization C of the closure of x' in X is a regular curve with the desired property. □

6 PROOF OF THE MAIN THEOREM

In this section we prove our main result. We say that "resolution of singularities holds for schemes of dimension $\leq d$ over k" if the following two conditions are satisfied.

(1) For any integral separated scheme of finite type X of dimension $\leq d$ over k, there exists a projective birational morphism $Y \to X$ with Y smooth over k which is an isomorphism over the regular locus of X.
(2) For any integral smooth scheme X of dimension $\leq d$ over k and any birational proper morphism $Y \to X$ there exists a tower of morphisms $X_n \to X_{n-1} \to \cdots \to X_0 = X$, such that $X_n \to X_{n-1}$ is a blow-up with a smooth center for $i = 1, \ldots, n$, and such that the composite morphism $X_n \to X$ factors through $Y \to X$.

THEOREM 6.1 (=THEOREM 1.1). *Let k be an algebraically closed field of characteristic $p \geq 0$, X a separated scheme of finite type over k and m a natural number. Then*

$$rec_X : H_1^S(X, \mathbb{Z}/m\mathbb{Z}) \longrightarrow \pi_1^{t,ab}(X)/m$$

is surjective. It is an isomorphism of finite abelian groups if $(m,p) = 1$, and for general m if resolution of singularities holds for schemes of dimension $\leq \dim X + 1$ over k.

The proof will occupy the rest of this section. Following the notation of Section 5, we write $H_t^0 = H_{\mathrm{et}}^0$ and consider the maps

$$rec_{i,X} : H_i^S(X, \mathbb{Z}/m\mathbb{Z}) \to H_t^i(X, \mathbb{Z}/m\mathbb{Z})^*$$

for $i = 0, 1$ (i.e., $rec_X = rec_{1,X}$). Given a morphism $X' \to X$, we have a commutative diagram of pairings defining rec_i for $i = 0, 1$.

$$\begin{array}{ccccc}
H_i^S(X', \mathbb{Z}/m\mathbb{Z}) & \times & H_t^i(X', \mathbb{Z}/m\mathbb{Z}) & \xrightarrow{\langle\, ,\, \rangle} & \mathbb{Z}/m\mathbb{Z} \\
\phi_* \big\downarrow & & \big\uparrow \phi^* & & \big\| \\
H_i^S(X, \mathbb{Z}/m\mathbb{Z}) & \times & H_t^i(X, \mathbb{Z}/m\mathbb{Z}) & \xrightarrow{\langle\, ,\, \rangle} & \mathbb{Z}/m\mathbb{Z}.
\end{array}$$

Step 1: $rec_{1,X}$ is surjective for arbitrary X.

We may assume that X is reduced and proceed by induction on $d = \dim X$. The case $\dim X = 0$ is trivial. Consider the normalization morphism $X' \to X$, which is an isomorphism outside a closed subscheme $Z \subset X$ of dimension $\leq d - 1$. Using the exact sequences of Propositions 5.1 and 5.2, which are compatible by Proposition 5.5 and the fact that $rec_{0,X}$ is an isomorphism by Lemma 5.3, a diagram chase shows that it suffices to show surjectivity of $rec_{1,X}$ for normal schemes.
Let X be normal. Since $H_t^1(X, \mathbb{Z}/m\mathbb{Z})$ is finite, it suffices to show that the pairing defining $rec_{1,X}$ has a trivial right kernel. We may assume that X is connected. Let $b \in H_t^1(X, \mathbb{Z}/m\mathbb{Z})$ be arbitrary but non-zero. By Proposition 5.6, we find a morphism $\phi : C \to X$ with C a smooth curve such that $\phi^*(b) \in H_{\mathrm{et}}^1(C, \mathbb{Z}/m\mathbb{Z})$ is non-zero. Since the pairing for C is perfect by Theorem 4.1, the pairing for X has a trivial right kernel.

Step 2: Theorem 6.1 holds if $(m,p) = 1$.

If $(m,p) = 1$, $H_1^S(X, \mathbb{Z}/m\mathbb{Z})$ and $H_{\mathrm{et}}^1(X, \mathbb{Z}/m\mathbb{Z})^*$ are isomorphic finite abelian groups by [SV1]. In particular, they have the same order. Hence the surjective homomorphism $rec_{1,X}$ is an isomorphism.

Step 3: Theorem 6.1 holds for arbitrary X if $m = p^r$ and resolution of singularities holds for schemes of dimension $\leq \dim X + 1$ over k.

We may assume that X is reduced. Using resolution of singularities and Chow's Lemma, we obtain a morphism $X' \to X$ with X' smooth and quasi-projective, which is an isomorphism over a dense open subscheme of X. Using the exact sequences of Propositions 5.1 and 5.2, Lemma 5.3, *Step 1*, induction on the dimension and the five-lemma, it suffices to show the result for smooth, quasi-projective schemes.
Let X be smooth, quasi-projective and let $\bar{X}$ be a smooth, projective variety containing X as a dense open subscheme. Then, by [Ge3, §5], we have an isomorphism

$$H^1_{\mathrm{et}}(\bar{X}, \mathbb{Z}/p^r\mathbb{Z})^* \cong \mathrm{CH}_0(\bar{X}, 1, \mathbb{Z}/p^r\mathbb{Z}).$$

Furthermore, by [SS, Thm. 2.7] (which makes the assumption $\dim \le 2$ but does not use it in its proof), we have an isomorphism

$$\mathrm{CH}_0(\bar{X}, 1, \mathbb{Z}/p^r\mathbb{Z}) \cong H^S_1(\bar{X}, \mathbb{Z}/p^r\mathbb{Z}).$$

By Proposition 6.2 below, the natural homomorphism

$$H^S_1(X, \mathbb{Z}/p^r\mathbb{Z}) \to H^S_1(\bar{X}, \mathbb{Z}/p^r\mathbb{Z})$$

is an isomorphism of finite abelian groups and by Proposition 2.10, we have an isomorphism

$$H^1_{\mathrm{et}}(\bar{X}, \mathbb{Z}/p^r\mathbb{Z}) \xrightarrow{\sim} H^1_t(X, \mathbb{Z}/p^r\mathbb{Z}).$$

Hence the finite abelian groups $H^1_t(X, \mathbb{Z}/p^r\mathbb{Z})^*$ and $H^S_1(X, \mathbb{Z}/p^r\mathbb{Z})$ are isomorphic, in particular, they have the same order. Since $rec_{1,X}$ is surjective, it is an isomorphism.

In order to conclude the proof of Theorem 6.1 it remains to show

PROPOSITION 6.2. *Let k be a perfect field, $X \in \mathrm{Sch}/k$ smooth, $U \subset X$ a dense open subscheme and $n \ge 0$ an integer. Assume that resolution of singularities holds for schemes of dimension $\le \dim X + n$ over k. Then for any $r \ge 1$ the natural map*

$$H^S_i(U, \mathbb{Z}/p^r\mathbb{Z}) \to H^S_i(X, \mathbb{Z}/p^r\mathbb{Z})$$

is an isomorphism of finite abelian groups for $i = 0, \ldots, n$.

REMARK 6.3. A proof of Proposition 6.2 for $n = 1$ and k algebraically closed independent of the assumption on resolution of singularities would relax the condition in Theorem 6.1 to:
There exists a smooth, projective scheme $\bar{X}' \in \mathrm{Sch}/k$, dense open subschemes $U' \subset X' \subset \bar{X}'$, $U \subset X$, and a surjective, proper morphism $X' \to X$ which induces an isomorphism $U'_{\mathrm{red}} \to U_{\mathrm{red}}$.
In particular, Theorem 6.1 would hold for $\dim X \le 3$ without any assumption on resolution of singularities [CV].

Proof of Proposition 6.2. We set $R = \mathbb{Z}/p^r\mathbb{Z}$. By [MVW], Lecture 14, we have

$$H_i^S(X,R) = \mathrm{Hom}_{\mathrm{DM}^{\mathrm{eff},-}_{\mathrm{Nis}}(k,R)}(R[i], M(X,R)).$$

Let $d = \dim X$. Choose a series of open subschemes $U = X_d \subset \cdots X_1 \subset X_0 = X$ such that $Z_j := X_j \smallsetminus X_{j+1}$ is smooth of dimension j for $j = 0, \ldots, d-1$. Using the exact Gysin triangles [MVW, 15.15]

$$M(X_{j+1},R) \to M(X_j,R) \to M(Z_j,R)(d-j)[2d-2j] \xrightarrow{[1]} M(X_{j+1},R)[1]$$

and induction, it suffices to show that

$$\mathrm{Hom}_{\mathrm{DM}^{\mathrm{eff},-}_{\mathrm{Nis}}(k,R)}(R[i], M(Z_j,R)(s)[2s]) = 0$$

for $j = 0, \ldots, d-1$, $i = 0, \ldots, n+1$ and $s \geq 1$. Using smooth compactifications of the Z_j and induction again, it suffices to show

$$\mathrm{Hom}_{\mathrm{DM}^{\mathrm{eff},-}_{\mathrm{Nis}}(k,R)}(R[i], M(Z,R)(s)[2s]) = 0$$

for Z connected, smooth, projective, $i = 0, \ldots, d - d_Z + n$ and $s \geq 1$.
By the comparison of higher Chow groups and motivic cohomology [V] and by [GL], Thm. 8.5, the restriction of $R(s)$ to the small Nisnevich site of a smooth scheme Y is isomorphic to $\nu_r^s[-s]$, where ν_r^s is the logarithmic de Rham Witt sheaf of Milne and Illusie. In particular, $R(s)|_Y$ is trivial for $s > \dim Y$.
For an étale k-scheme Z we obtain

$$\mathrm{Hom}_{\mathrm{DM}^{\mathrm{eff},-}_{\mathrm{Nis}}(k,R)}(R[i], M(Z,R)(s)[2s]) = H^{2s-i}_{\mathrm{Nis}}(Z, R(s)) = 0$$

for $s \geq 1$ and all $i \geq 0$. Now assume $\dim Z \geq 1$. Using resolution of singularities for schemes of dimension $\leq d+n$, the same method as in the proof of [SS], Thm. 2.7 yields isomorphisms

$$\mathrm{Hom}_{\mathrm{DM}^{\mathrm{eff},-}_{\mathrm{Nis}}(k,R)}(R[i], M(Z,R)) \cong \mathrm{CH}^{d_Z}(Z,i,R)$$

for $i = 0, \ldots, d-1+n$. Applying this to $Z \times \mathbb{P}^s$ and using the decompositions given by the projective bundle theorem on both sides implies isomorphisms

$$\mathrm{Hom}_{\mathrm{DM}^{\mathrm{eff},-}_{\mathrm{Nis}}(k,R)}(R[i], M(Z,R)(s)[2s]) \cong \mathrm{CH}^{d_Z+s}(Z,i,R)$$

for $i = 0, \ldots, d-1+n$. By [V], the latter group is isomorphic to

$$\mathrm{Hom}_{\mathrm{DM}^{\mathrm{eff},-}_{\mathrm{Nis}}(k,R)}(M(Z,R)[2d_Z+2s-i], R(d_Z+s)) \cong H^{2d_Z+2s-i}_{\mathrm{Nis}}(Z, R(d_Z+s)),$$

which vanishes for $s \geq 1$. This finishes the proof. □

REMARK 6.4. The assertion of Proposition 6.2 remains true for non-smooth X if U contains the singular locus of X (see [Ge4], Prop. 3.3).

7 COMPARISON WITH THE ISOMORPHISM OF SUSLIN-VOEVODSKY

THEOREM 7.1. *Let k be an algebraically closed field, $X \in \mathrm{Sch}/k$ and m an integer prime to $\mathrm{char}(k)$. Then the reciprocity isomorphism*

$$rec_X : H_1^S(X, \mathbb{Z}/m\mathbb{Z}) \longrightarrow \pi_1^{ab}(X)/m$$

is the dual of the isomorphism

$$\alpha_X : H^1_{\mathrm{et}}(X, \mathbb{Z}/m\mathbb{Z}) \longrightarrow H^1_S(X, \mathbb{Z}/m\mathbb{Z})$$

of [SV1], Cor. 7.8.

The proof will occupy the rest of this section. Let $i : \mathbb{Z}/m\mathbb{Z} \hookrightarrow I^0$ be an injection into an injective sheaf in the category of $\mathbb{Z}/m\mathbb{Z}$-module sheaves on $(\mathrm{Sch}/k)_{\mathrm{qfh}}$ and put $J^1 = \mathrm{coker}(i)$. Then (see the end of section 3) the pairing between $H_1^S(X, \mathbb{Z}/m\mathbb{Z})$ and $H^1_{\mathrm{et}}(X, \mathbb{Z}/m\mathbb{Z})$ constructed in Proposition 2.12 can be given as follows: For $a \in H_1^S(X, \mathbb{Z}/m\mathbb{Z})$ choose a representing correspondence $\alpha \in \mathrm{Cor}(\Delta^1, X)$ and for $b \in H^1_{\mathrm{et}}(X, \mathbb{Z}/m\mathbb{Z})$ a pre-image $\beta \in J^1(X)$. Consider the diagram

$$\begin{tikzcd} & I^0(X) \arrow[r] \arrow[d, "\alpha^*"] & J^1(X) \arrow[d, "\alpha^*"] \\ & I^0(\Delta^1) \arrow[r, two heads] \arrow[d, "0^* - 1^*"] & J^1(\Delta^1) \arrow[d, "0^* - 1^*"] \\ \mathbb{Z}/m\mathbb{Z} \arrow[r, hook] & I^0(\Delta^0) \arrow[r] & J^0(\Delta^0). \end{tikzcd} \tag{8}$$

Then $\alpha^*(\beta)$ is the image of some element $\gamma \in I^0(\Delta^1)$ and $(0^* - 1^*)(\gamma) \in \mathbb{Z}/m\mathbb{Z} = \ker(I^0(\Delta^0) \to J^1(\Delta^0))$ equals $\langle a, b \rangle$.

For $Y \in \mathrm{Sch}/k$ let $\mathbb{Z}_Y^{\mathrm{qfh}}$ be the free qfh-sheaf generated by Y. We set $A = \mathbb{Z}[1/\mathrm{char}(k)]$ and $L_Y = \mathbb{Z}_Y^{\mathrm{qfh}} \otimes A$. For smooth U the homomorphism

$$\mathrm{Cor}(U, X) \otimes A \to \mathrm{Hom}_{\mathrm{qfh}}(L_U, L_X)$$

is an isomorphism by [SV1], Thm. 6.7. We have

$$H^1_{\mathrm{et}}(X, \mathbb{Z}/m\mathbb{Z}) = H^1_{\mathrm{qfh}}(X, \mathbb{Z}/m\mathbb{Z}) = \mathrm{Ext}^1_{\mathrm{qfh}}(L_X, \mathbb{Z}/m\mathbb{Z})$$

$$= \mathrm{coker}(\mathrm{Hom}_{\mathrm{qfh}}(L_X, I^0) \to \mathrm{Hom}_{\mathrm{qfh}}(L_X, J^1)).$$

The diagram (8) can be rewritten in terms of Hom-groups as follows:

$$\begin{tikzcd}
& \mathrm{Hom}_{\mathrm{qfh}}(L_X, I^0) \arrow[r] \arrow[d, "\alpha^*"] & \mathrm{Hom}_{\mathrm{qfh}}(L_X, J^1) \arrow[d, "\alpha^*"] \\
& \mathrm{Hom}_{\mathrm{qfh}}(L_{\Delta^1}, I^0) \arrow[r, two heads] \arrow[d, "0^* - 1^*"] & \mathrm{Hom}_{\mathrm{qfh}}(L_{\Delta^1}, J^1) \arrow[d, "0^* - 1*"] \\
\mathbb{Z}/m\mathbb{Z} \arrow[r, hook] & \mathrm{Hom}_{\mathrm{qfh}}(L_{\Delta^0}, I^0) \arrow[r] & \mathrm{Hom}_{\mathrm{qfh}}(L_{\Delta^0}, J^1).
\end{tikzcd} \tag{9}$$

We denote the morphism $L_X \to J^1$ corresponding to $\beta \in J^1(X) \cong \mathrm{Hom}_{\mathrm{qfh}}(L_X, J^1)$ by the same letter β. Putting $E := I^0 \times_{J_1,\beta} L_X$, the extension

$$0 \longrightarrow \mathbb{Z}/m\mathbb{Z} \longrightarrow E \longrightarrow L_X \longrightarrow 0$$

represents $b \in \mathrm{Ext}^1_{\mathrm{qfh}}(L_X, \mathbb{Z}/m\mathbb{Z})$. Consider the diagram

$$\begin{tikzcd}
& \mathrm{Hom}_{\mathrm{qfh}}(L_X, E) \arrow[r] \arrow[d, "\alpha^*"] & \mathrm{Hom}_{\mathrm{qfh}}(L_X, L_X) \arrow[d, "\alpha^*"] \\
& \mathrm{Hom}_{\mathrm{qfh}}(L_{\Delta^1}, E) \arrow[r, two heads] \arrow[d, "0^* - 1^*"] & \mathrm{Hom}_{\mathrm{qfh}}(L_{\Delta^1}, L_X) \arrow[d, "0^* - 1^*"] \\
\mathbb{Z}/m\mathbb{Z} \arrow[r, hook] & \mathrm{Hom}_{\mathrm{qfh}}(L_{\Delta^0}, E) \arrow[r] & \mathrm{Hom}_{\mathrm{qfh}}(L_{\Delta^0}, L_X).
\end{tikzcd} \tag{10}$$

Because diagram (10) maps to diagram (9) via β_* and $\mathrm{id} \in \mathrm{Hom}_{\mathrm{qfh}}(L_X, L_X)$ maps under β_* to $\beta \in \mathrm{Hom}_{\mathrm{qfh}}(L_X, J^1)$, we can calculate the pairing using diagram (10) after replacing β by id. Since id maps to $\alpha \in \mathrm{Hom}_{\mathrm{qfh}}(L_{\Delta^1}, L_X)$ under α^*, we see, writing the lower part of diagram (10) in the form

$$\begin{tikzcd}
\mathbb{Z}/m\mathbb{Z} \arrow[r, hook] \arrow[d, "0"'] & E(\Delta^1) \arrow[r, two heads] \arrow[d, "0^* - 1^*"'] & L_X(\Delta^1) \arrow[d, "0^* - 1^*"] \arrow[dl, dashed, "h"] \\
\mathbb{Z}/m\mathbb{Z} \arrow[r, hook] & E(\Delta^0) \arrow[r] & L_X(\Delta^0),
\end{tikzcd} \tag{11}$$

that

$$\langle a, b \rangle = h(\alpha) \bmod m \in \ker(E(\Delta^0)/m \to L_X(\Delta^0)/m) = \mathbb{Z}/m\mathbb{Z},$$

where h is the unique homomorphism making diagram (11) commutative. We consider the complex $C_\bullet(X) = \mathrm{Cor}(\Delta^\bullet, X) \otimes A = L_X(\Delta^\bullet)$ with the obvious differentials. By the above considerations, the homomorphism induced by the pairing of Proposition 2.12

$$H^1_{\mathrm{et}}(X, \mathbb{Z}/m\mathbb{Z}) = H^1_{\mathrm{qfh}}(X, \mathbb{Z}/m\mathbb{Z}) \longrightarrow$$
$$H^S_1(X, \mathbb{Z}/m\mathbb{Z})^* = \mathrm{Ext}^1_A(C_\bullet(X), \mathbb{Z}/m\mathbb{Z}) = \mathrm{Hom}_{D(A)}(C_\bullet(X), \mathbb{Z}/m\mathbb{Z}[1]),$$

is given by sending an extension class $[\mathbb{Z}/m\mathbb{Z} \hookrightarrow E \twoheadrightarrow L_X]$ to the morphism $C_\bullet(X) \to \mathbb{Z}/m\mathbb{Z}[1]$ in the derived category of A-modules represented by the morphism

$$C_\bullet(X) \to [0 \to E(\Delta^0) \to L_X(\Delta^0) \to 0]$$

which is given by $\mathrm{id} : L_X(\Delta^0) \to L_X(\Delta^0)$ in degree zero and by $h : L_X(\Delta^1) \to E(\Delta^0)$ in degree one.
The same construction works for any qfh-sheaf of A-modules F instead of L_X, i.e., setting $C_\bullet(F) = F(\Delta^\bullet)$ and starting from an element

$$[\mathbb{Z}/m\mathbb{Z} \hookrightarrow E \twoheadrightarrow F] \in \mathrm{Ext}^1_{\mathrm{qfh}}(F, \mathbb{Z}/m\mathbb{Z}),$$

we get a map $C_\bullet(F) \to \mathbb{Z}/m\mathbb{Z}[1]$ in the derived category of A-modules. We thus constructed a homomorphism

$$\mathrm{Ext}^1_{\mathrm{qfh}}(F, \mathbb{Z}/m\mathbb{Z}) \longrightarrow \mathrm{Ext}^1(C_\bullet(F), \mathbb{Z}/m\mathbb{Z}), \tag{12}$$

which for $F = L_X$ and under the canonical identifications coincides with the map

$$H^1_{\mathrm{et}}(X, \mathbb{Z}/m\mathbb{Z}) \longrightarrow H^S_1(X, \mathbb{Z}/m\mathbb{Z})^*$$

induced by the pairing constructed in Proposition 2.12.

Now we compare the map (12) with the map

$$\alpha_X : \mathrm{Ext}^1_{\mathrm{qfh}}(F, \mathbb{Z}/m\mathbb{Z}) \longrightarrow \mathrm{Ext}^1_A(C_\bullet(F), \mathbb{Z}/m\mathbb{Z}) \tag{13}$$

constructed by Suslin-Voevodsky [SV1] (cf. [Ge1] for the case of positive characteristic). Let $F^\sim_\bullet$ be the complex of qfh-sheaves associated with the complex of presheaves $F_\bullet(U) = F(U \times \Delta^\bullet)$. By [SV1], the inclusion $F \to F^\sim_\bullet$ induces an isomorphism

$$\mathrm{Ext}^1_{\mathrm{qfh}}(F^\sim_\bullet, \mathbb{Z}/m\mathbb{Z}) \xrightarrow{\sim} \mathrm{Ext}^1_{\mathrm{qfh}}(F, \mathbb{Z}/m\mathbb{Z}), \tag{14}$$

and evaluation at $Spec(k)$ induces an isomorphism

$$\mathrm{Ext}^1_{\mathrm{qfh}}(F^\sim_\bullet, \mathbb{Z}/m\mathbb{Z}) \xrightarrow{\sim} \mathrm{Ext}^1_A(C_\bullet(F), \mathbb{Z}/m\mathbb{Z}). \tag{15}$$

The map (13) of Suslin-Voevodsky is the composite of the inverse of (14) with (15).

We construct the inverse of (14). Let a class $[\mathbb{Z}/m\mathbb{Z} \hookrightarrow E \twoheadrightarrow F] \in \mathrm{Ext}^1_{\mathrm{qfh}}(F, \mathbb{Z}/m\mathbb{Z})$ be given. As a morphism in the derived category this class is given by the homomorphism

$$\begin{array}{ccccccc}
0 & \longrightarrow & 0 & \longrightarrow & F & \longrightarrow & 0 \\
\downarrow & & \downarrow & & \downarrow{\scriptstyle id} & & \downarrow \\
0 & \longrightarrow & E & \longrightarrow & F & \longrightarrow & 0
\end{array} \quad .$$

We therefore have to construct a homomorphism $F_1 \longrightarrow E$ making the diagram

$$\begin{array}{ccccccc}
F_2 & \longrightarrow & F_1 & \longrightarrow & F_0 & \longrightarrow & 0 \\
\downarrow & & \downarrow & & \downarrow{\scriptstyle id} & & \downarrow \\
0 & \longrightarrow & E & \longrightarrow & F & \longrightarrow & 0
\end{array}$$

commutative. The construction is a sheafified version of what we did before. Let $U \in \mathrm{Sch}/k$ be arbitrary. Consider the diagram

$$\begin{array}{ccccccccc}
0 & \longrightarrow & \mathbb{Z}/m\mathbb{Z}(U) & \longrightarrow & E(U \times \Delta^2) & \longrightarrow & F(U \times \Delta^2) & \longrightarrow & 0 \\
& & \downarrow{\scriptstyle id} & & \downarrow{\scriptstyle \delta^0 - \delta^1 + \delta^2} & & \downarrow{\scriptstyle \delta^0 - \delta^1 + \delta^2} & & \\
0 & \longrightarrow & \mathbb{Z}/m\mathbb{Z}(U) & \longrightarrow & E(U \times \Delta^1) & \longrightarrow & F(U \times \Delta^1) & \longrightarrow & 0 \\
& & \downarrow{\scriptstyle 0} & & \downarrow{\scriptstyle 0^* - 1^*} & & \downarrow{\scriptstyle 0^* - 1^*} & & \\
0 & \longrightarrow & \mathbb{Z}/m\mathbb{Z}(U) & \longrightarrow & E(U) & \longrightarrow & F(U) & \longrightarrow & 0
\end{array} \quad .$$

Let $\alpha_1 \in F(U \times \Delta^1)$ be given. By the smooth base change theorem and since $H^1_{\mathrm{et}}(\Delta^1, \mathbb{Z}/m\mathbb{Z}) = 0$, we can lift α_1 to $E(U \times \Delta^1)$ after replacing U by a sufficiently fine étale cover. Applying $0^* - 1^*$ to this lift, we get an element in $E(U)$. This gives the homomorphism $F_1 \to E$. Now let $\alpha_2 \in F(U \times \Delta^2)$ be arbitrary. After replacing U by a sufficiently fine étale cover, we can lift α_2 to $E(U \times \Delta^2)$. Since $(0^* - 1^*)(\delta^0 - \delta^1 + \delta^2) = 0$ this shows that $(\delta^0 - \delta^1 + \delta^2)(\alpha_2)$ maps to zero in $E(U)$. This describes the inverse isomorphism to (14). Evaluating at $U = Spec(k)$ gives back our original construction, hence (12) and (13) are the same maps. This finishes the proof.

REFERENCES

[CV] V. Cossart, O. Piltant *Resolution of singularities of threefolds in positive characteristic. II*. J. Algebra 321 (2009), no. 7, 1836–1976.

[Fr] E. M. Friedlander *Etale homotopy of simplicial schemes*. Ann. of Math. Studies 104 Princeton Univ. Press 1982

[Ge1] T. Geisser *Applications of de Jong's theorem on alterations*. In H. Hauser, J. Lipman, F. Oort, A. Quirós (eds.): Resolution of Singularities. In tribute of Oscar Zariski. Progress in Math., Birkhäuser 1999

[Ge2] T. Geisser *Arithmetic cohomology over finite fields and special values of ζ-functions*. Duke Math. J. 133 (2006), no. 1, 27–57.

[Ge3] T. Geisser *Duality via cycle complexes*. Ann. of Math. (2) 172 (2010), no. 2, 1095–1126.

[Ge4] T. Geisser *On Suslin's singular homology and cohomology*. Doc. Math. 2010, Extra volume: Andrei A. Suslin's sixtieth birthday, 223–249.

[GL] T. Geisser, M. Levine *The K-theory of fields in characteristic p*. Inv. Math. 139 (2000), 459–494.

[FJ] M. D. Fried, M. Jarden *Field arithmetic*. Ergebnisse der Mathematik und ihrer Grenzgebiete. (3), 11. Springer-Verlag, Berlin, 1986.

[KS] M. Kerz, A. Schmidt *On different notions of tameness in arithmetic geometry*. Math. Ann. 346 (2010), 641–668.

[Li] S. Lichtenbaum *Suslin homology and Deligne 1-motives*. in P. G. Goerss, J. F. Jardine (ed.): Algebraic K-theory and Algebraic Topology 1993, 189–196.

[MVW] C. Mazza, V. Voevodsky and Ch. Weibel *Lectures on Motivic Cohomology*. Clay Monographs in Math. 2, AMS 2006.

[Mi] J. S. Milne *Étale Cohomology*. Princeton University Press 1980.

[Mu] D. Mumford *Abelian varieties*. Tata Institute of Fundamental Research Studies in Mathematics, No. 5, Oxford University Press, London 1970.

[SS] A. Schmidt, M. Spieß *Singular homology and class field theory of varieties over finite fields*. J. reine angew. Math. 527 (2000) 13–37.

[Se] J.-P. Serre *Groupes algébriques et corps de classes*. Deuxième édition. Hermann, Paris, 1975.

[SV1] A. Suslin, V. Voevodsky *Singular homology of abstract algebraic varieties*. Invent. Math. 123 (1996), 61–94.

[SV2] A. A. Suslin, V. Voevodsky *Bloch-Kato conjecture and motivic cohomology with finite coefficients.* The arithmetic and geometry of algebraic cycles (Banff, AB, 1998), 117-189, NATO Sci. Ser. C Math. Phys. Sci., 548, Kluwer Acad. Publ., Dordrecht, 2000.

[V] V. Voevodsky *Motivic cohomology groups are isomorphic to higher Chow groups in any characteristic*. Int. Math. Res. Not. 2002, no. 7, 351–355.

Thomas Geisser
Rikkyo University
Department of Mathematics
3-34-1 Nishi-Ikebukuro
Toshima-ku
Tokyo Japan 171-8501
geisser@rikkyo.ac.jp

Alexander Schmidt
Universität Heidelberg
Mathematisches Institut
Im Neuenheimer Feld 288
D-69120 Heidelberg
Deutschland
schmidt@mathi.uni-heidelberg.de

Lifting from two Elliptic Modular Forms to Siegel Modular Forms of Half-Integral Weight of even Degree

Shuichi Hayashida

Received: February 24, 2015
Revised: December 12, 2015

Communicated by Don Blasius

Abstract. The aim of this paper is to construct lifts from two elliptic modular forms to Siegel modular forms of half-integral weight of even degree under the assumption that the constructed Siegel modular form is not identically zero. The key ingredient of the proof is a new Maass relation for Siegel modular forms of half-integral weight and any degree.

2010 Mathematics Subject Classification: 11F46 (primary), 11F37, 11F50 (secondary)
Keywords and Phrases: Siegel modular forms, Jacobi forms, Maass relation

1 Introduction

1.1

Lifts from two elliptic modular forms to Siegel modular form of half-integral weight of degree two have been conjectured by Ibukiyama and the author[H-I 05]. In the present article we will give a partial answer for the conjecture in [H-I 05] and shall generalize these lifts as lifts from two elliptic modular forms to Siegel modular forms of half-integral weight of *any even degree* (Theorem 8.3).
The construction of the lift can be regarded as a half-integral weight version of the Miyawaki-Ikeda lift. The Miyawaki-Ikeda lift has been shown by Ikeda [Ik 06]. In the present article we will give a proof to the fact that constructed Siegel modular forms of half-integral weight are eigenforms, if it does not identically vanish. Moreover, we will compute the L-function of the constructed Siegel modular forms of half-integral weight. The key ingredient of

the proof of the lift in the present article is to introduce a generalized Maass relation for Siegel modular forms of half-integral weight (Theorem 7.6, 8.2). Generalized Maass relations are relations among Fourier-Jacobi coefficients of Siegel modular forms and are regarded as relations among Fourier coefficients. Theorem 7.6 is a generalization of the Maass relation for generalized Cohen-Eisenstein series, which is a Siegel modular form of half-integral weight of general degree. And Theorem 8.2 is a generalization of the Maass relation for Siegel cusp forms of half-integral weight of odd degree.

1.2

We explain our results more precisely.
We denote by $M^+_{k-\frac{1}{2}}(\Gamma_0^{(n)}(4))$ the generalized plus-space of weight $k-\frac{1}{2}$ of degree n, which is a subspace of Siegel modular forms of half-integral weight and is a generalization of the Kohnen plus-space (see [Ib 92] or §4.3 for the definition of generalized plus-space). Let $F \in M^+_{k-\frac{1}{2}}(\Gamma_0^{(n)}(4))$ be an eigenform for any Hecke operators. We put

$$Q_{F,p}(z) \quad = \quad \prod_{i=0}^{n}(1-\mu_{i,p}z)(1-\mu_{i,p}^{-1}z),$$

where complex numbers $\{\mu_{i,p}^{\pm}\}$ are p-parameters of F introduced in [Zh 84] if p is an odd prime. If $p=2$, then we define $\{\mu_{i,2}^{\pm}\}$ by using the isomorphism between generalized plus-space and the space of Jacobi forms of index 1. We denote the modified Zhuravlev L-function by

$$L(s,F) \quad := \quad \prod_{p} Q_{F,p}(p^{-s+k-\frac{3}{2}}).$$

The Zhuravlev L-function is originally introduced in [Zh 84] without the Euler 2-factor, which is a generalization of the L-function of elliptic modular forms of half-integral weight introduced in [Sh 73].
We denote by $S^+_{k-\frac{1}{2}}(\Gamma_0^{(n)}(4))$ the space of Siegel cusp forms in $M^+_{k-\frac{1}{2}}(\Gamma_0^{(n)}(4))$. The following theorem is the main result of this article.

Theorem 8.3. *Let k be an even integer and n be an integer greater than* 1*. Let $h \in S^+_{k-n+\frac{1}{2}}(\Gamma_0^{(1)}(4))$ and $g \in S^+_{k-\frac{1}{2}}(\Gamma_0^{(1)}(4))$ be eigenforms for all Hecke operators. Then there exists a $\mathcal{F}_{h,g} \in S^+_{k-\frac{1}{2}}(\Gamma_0^{(2n-2)})$. Under the assumption that $\mathcal{F}_{h,g}$ is not identically zero, then $\mathcal{F}_{h,g}$ is an eigenform with the L-function which satisfies*

$$L(s,\mathcal{F}_{h,g}) \quad = \quad L(s,g)\prod_{i=1}^{2n-3} L(s-i,h).$$

By numerical computations of Fourier coefficients of $\mathcal{F}_{h,g}$ we checked that $\mathcal{F}_{h,g}$ does not identically vanish for some (n,k). (See §9 for the detail).

Remark that the above theorem was first conjectured by Ibukiyama and the author [H-I 05] in the case of $n = 2$ not only for even integer k, but also for odd integer k.

The construction of $\mathcal{F}_{h,g}$ was suggested by T. Ikeda to the author, which is given by a composition of three maps and an inner product. These three maps are a Ikeda lift (Duke-Imamoglu-Ibukiyama-Ikeda lift), a map of the Fourier-Jacobi expansion and an isomorphism between Jacobi forms of index 1 and Siegel modular forms of half-integral weight. In §8 we will explain the detail of the construction of $\mathcal{F}_{h,g}$.

To prove Theorem 8.3 we use a generalized Maass relation for generalized Cohen-Eisenstein series (Theorem 7.6). Once we obtain Theorem 7.6, it is not so hard to show Theorem 8.3. The most part of this article is devoted to show Theorem 7.6. We now explain the generalized Maass relation for generalized Cohen-Eisenstein series (Theorem 7.6).

Let k be an *even* integer and $\mathcal{H}^{(n+1)}_{k-\frac{1}{2}}$ be the generalized Cohen-Eisenstein series of degree $n+1$ of weight $k-\frac{1}{2}$ (see §4.4 for the definition of generalized Cohen-Eisenstein series). The form $\mathcal{H}^{(n+1)}_{k-\frac{1}{2}}$ is a Siegel modular form of weight $k-\frac{1}{2}$ of degree $n+1$.

For integer m, we denote by $e^{(n)}_{k-\frac{1}{2},m}$ the m-th Fourier-Jacobi coefficient of $\mathcal{H}^{(n+1)}_{k-\frac{1}{2}}$:

$$\mathcal{H}^{(n+1)}_{k-\frac{1}{2}}\left(\begin{pmatrix} \tau & z \\ {}^t z & \omega \end{pmatrix}\right) = \sum_{\substack{m\geq 0 \\ m\equiv 0,3 \bmod 4}} e^{(n)}_{k-\frac{1}{2},m}(\tau,z)\, e^{2\pi\sqrt{-1}m\omega}, \tag{1.1}$$

where $\tau \in \mathfrak{H}_n$ and $\omega \in \mathfrak{H}_1$, and where $\mathfrak{H}_n$ denotes the Siegel upper half space of degree n. We denote by $J^{(n)}_{k-\frac{1}{2},m}$ the space of Jacobi forms of degree n of weight $k-\frac{1}{2}$ of index m (cf. §2.6) and denote by $J^{(n)*}_{k-\frac{1}{2},m}$ (cf. §4.4) a subspace of $J^{(n)}_{k-\frac{1}{2},m}$. Then, the above form $e^{(n)}_{k-\frac{1}{2},m}$ belongs to $J^{(n)*}_{k-\frac{1}{2},m}$. Because $\mathcal{H}^{(n+1)}_{k-\frac{1}{2}}$ belongs to the generalized plus-space $M^+_{k-\frac{1}{2}}(\Gamma_0^{(n+1)}(4))$, we can show that the form $e^{(n)}_{k-\frac{1}{2},m}$ is identically zero unless $m \equiv 0$, $3 \bmod 4$.

We denote by $M_k(\Gamma_{n+2})$ the space of Siegel modular forms of weight k of degree $n+2$ and denote by $J^{(n+1)}_{k,1}$ the space of Jacobi forms of weight k of index 1 of degree $n+1$. We denote by $E^{(n)}_k \in M_k(\Gamma_n)$ the Siegel-Eisenstein series of weight k of degree n (cf. (3.2) in §3) and by $E^{(n)}_{k,1} \in J^{(n)}_{k,1}$ the Jacobi-Eisenstein series of weight k of index 1 of degree n (cf. (3.1) in §3). The form $\mathcal{H}^{(n+1)}_{k-\frac{1}{2}}$ is constructed from $E^{(n+1)}_{k,1}$. The diagram of the above correspondence is

$$\begin{tikzcd}
E_k^{(n+2)} \in M_k(\Gamma_{n+2}) \arrow[d] & \\
E_{k,1}^{(n+1)} \in J_{k,1}^{(n+1)} \arrow[r] & \mathcal{H}_{k-\frac{1}{2}}^{(n+1)} \in M_{k-\frac{1}{2}}^{+}(\Gamma_0^{(n+1)}(4)) \arrow[d] \\
& \left\{e_{k-\frac{1}{2},m}^{(n)}\right\}_m \in \bigotimes_{m\equiv 0,3 \bmod 4} J_{k-\frac{1}{2},m}^{(n)*}.
\end{tikzcd}$$

In §2.7 (for any odd prime p) and in §4.7 (for $p=2$) we will introduce index-shift maps $\tilde{V}_{\alpha,n-\alpha}(p^2)$ ($\alpha = 0,...,n$), which are linear maps from $J_{k-\frac{1}{2},m}^{(n)*}$ to the space of holomorphic functions on $\mathfrak{H}_n \times \mathbb{C}^{(n,1)}$. If p is odd then $\tilde{V}_{\alpha,n-\alpha}(p^2)$ is a linear map from $J_{k-\frac{1}{2},m}^{(n)*}$ to $J_{k-\frac{1}{2},mp^2}^{(n)}$. These maps $\tilde{V}_{\alpha,n-\alpha}(p^2)$ are generalizations of the V_l-map in [E-Z 85, p.43] for half-integral weight of general degrees. For any $\phi \in J_{k-\frac{1}{2},m}^{(n)}$ and for any integer a we define $(\phi|U_a)(\tau,z) := \phi(\tau,az)$.
The following theorem is a generalization of the Maass relation for the generalized Cohen-Eisenstein series, where we use the symbol

$$\begin{aligned}
&e_{k-\frac{1}{2},m}^{(n)}|(\tilde{V}_{0,n}(p^2),\tilde{V}_{1,n-1}(p^2),...,\tilde{V}_{n,0}(p^2))\\
&:= (e_{k-\frac{1}{2},m}^{(n)}|\tilde{V}_{0,n}(p^2), e_{k-\frac{1}{2},m}^{(n)}|\tilde{V}_{1,n-1}(p^2),...,e_{k-\frac{1}{2},m}^{(n)}|\tilde{V}_{n,0}(p^2)).
\end{aligned}$$

Theorem 7.6. *Let $e_{k-\frac{1}{2},m}^{(n)}$ be the m-th Fourier-Jacobi coefficient of generalized Cohen-Eisenstein series $H_{k-\frac{1}{2}}^{(n+1)}$. (See (1.1)). Then we obtain*

$$\begin{aligned}
&e_{k-\frac{1}{2},m}^{(n)}|(\tilde{V}_{0,n}(p^2),\tilde{V}_{1,n-1}(p^2),...,\tilde{V}_{n,0}(p^2))\\
=\ & p^{k(n-1)-\frac{1}{2}(n^2+5n-5)}\left(e_{k-\frac{1}{2},\frac{m}{p^2}}^{(n)}|U_{p^2},\, e_{k-\frac{1}{2},m}^{(n)}|U_p,\, e_{k-\frac{1}{2},mp^2}^{(n)}\right)\\
&\times \begin{pmatrix} 0 & p^{2k-3} \\ p^{k-2} & p^{k-2}\left(\frac{-m}{p}\right) \\ 0 & 1 \end{pmatrix} A_{2,n+1}^p\left(p^{k-\frac{n+2}{2}-\frac{1}{2}}\right) diag(1,p^{1/2},\cdots,p^{n/2}).
\end{aligned}$$

Here $A_{2,n+1}^p\left(p^{k-\frac{n+2}{2}-\frac{1}{2}}\right)$ is a $2\times(n+1)$ matrix which is introduced in the beginning of §7 and the both sides of the above identity are vectors of forms. For any prime p we regard $e_{k-\frac{1}{2},\frac{m}{p^2}}^{(n)}$ as zero, if $\frac{m}{p^2}$ is not an integer or $\frac{m}{p^2} \not\equiv 0$, $3 \mod 4$. *The symbol $\left(\frac{*}{p}\right)$ denotes the Legendre symbol for odd prime p, and $\left(\frac{a}{2}\right) := 0,1,-1$ accordingly as a is even, $a\equiv \pm 1$ mod 8 or $a \equiv \pm 3$ mod 8.*

Theorem 7.6 gives also a relation among Fourier coefficients of Siegel-Eisenstein series of integral weight. The Fourier coefficients of Ikeda lifts satisfy similar relations to the ones of the Fourier coefficients of Siegel-Eisenstein series (see Theorem 8.2 for the detail). We call these relations of Fourier coefficients of Ikeda lifts also the generalized Maass relations. The generalized Maass relation among Fourier coefficients of the Ikeda lift $I_{2n}(h)$ of h gives a fact that $\mathcal{F}_{h,g}$ in Theorem 8.3 is an eigenform for all Hecke operators, since the form $\mathcal{F}_{h,g}$ is constructed from $I_{2n}(h)$ (and g). Moreover, the eigenvalues of $\mathcal{F}_{h,g}$ are calculated from the generalized Maass relations of Fourier coefficients of $I_{2n}(h)$. This is the reason why we need Theorem 7.6 to show Theorem 8.3. For the detail of the proof of Theorem 8.3 see §8.

1.3 About generalized Cohen-Eisenstein series

We remark that the generalized Cohen-Eisenstein series has been introduced by Arakawa [Ar 98]. These series are Siegel modular forms of half-integral weight. The Cohen-Eisenstein series were originally introduced by Cohen [Co 75] as one variable functions. In the case of degree one, it is known that the Cohen-Eisenstein series correspond to the Eisenstein series with respect to $\mathrm{SL}(2,\mathbb{Z})$ by the Shimura correspondence. The generalized Cohen-Eisenstein series is defined from the Jacobi-Eisenstein series of index 1 through the isomorphism between Jacobi forms of index 1 and Siegel modular forms of half-integral weight.

1.4 About generalized Maass relations

As for generalizations of the Maass relation, Yamazaki [Yk 86, Yk 89] obtained some relations among Fourier-Jacobi coefficients of Siegel-Eisenstein series of arbitrary degree of integral weight of *integer indices*. For our purpose we generalize some results in [Yk 86, Yk 89] on Fourier-Jacobi coefficients of Siegel-Eisenstein series of integer indices to indices of half-integral symmetric matrix of size 2. Here the right-lower part or the left-upper part of these matrices of the index is 1. We need to introduce index-shift maps on Jacobi forms of indices of such matrix (cf. §2.7). To calculate the action of index-shift maps on Fourier-Jacobi coefficients of Siegel-Eisenstein series, we use a relation between Fourier-Jacobi coefficients of Siegel-Eisenstein series and Jacobi-Eisenstein series (cf. Proposition 3.3). This relation has been shown by Boecherer [Bo 83, Satz7]. We also need to show a identity relation between Jacobi forms of integral weight of 2×2 matrix index and Jacobi forms of half-integral weight of integer index (Lemma 4.2). Moreover, we need to show a compatibility between this identity relation and index-shift maps (cf. Proposition 4.3, 4.4).

Through these relations we can show that the generalized Maass relation of generalized Cohen-Eisenstein series (Theorem 7.6) are equivalent to relations among Jacobi-Eisenstein series of integral weight of indices of matrix of size 2 (Proposition 7.4). Finally, to obtain the generalized Maass relation in Theorem 7.6, we need to calculate the action of index-shift maps on Jacobi-Eisenstein series of integral weight of indices of matrix of size 2 (cf. §5).

Remark 1.1
In his paper [Ko 02] Kohnen gives a generalization of the Maass relation for Siegel modular forms of even degree $2n$. His result is different from our generalization, since his result is concerned with the Fourier-Jacobi coefficients with $(2n-1)\times(2n-1)$ matrix index. We remark that some characterizations of the Ikeda lift by using generalized Maass relation in [Ko 02] are obtained by Kohnen-Kojima [KK 05] and by Yamana [Yn 10]. The characterization of the Ikeda lift by using the generalized Maass relation in Theorem 8.2 is open problem.

Remark 1.2
In his paper [Ta 86, §5] Tanigawa has obtained the same identity in Theorem 7.6 for *Siegel-Eisenstein series of half-integral weight of degree two* with arbitrary level N which satisfies $4|N$. He showed the identity by using the formula of local densities under the assumption $p \nmid N$. In our case we treat the *generalized Cohen-Eisenstein series* of arbitrary degree, which has essentially level 1. Hence our result contains the relation also for $p = 2$. Moreover, our result is valid for any general degree.

Remark 1.3
To show the generalized Maass relations in Theorem 7.6, 8.2, we treat the following four things:

1. Fourier-Jacobi expansion of Jacobi forms (cf. §4.1),
2. Fourier-Jacobi expansion of Siegel modular forms of half-integral weight (cf. §4.2),
3. An isomorphism between Jacobi forms of matrix index of integral weight and Jacobi forms of integer index of half-integral weight (cf. §4.5)
4. Exchange relations between the Siegel Φ-operator for Jacobi forms and the index-shift map for Jacobi forms of matrix index or of half-integral weight (cf. §6). This is an analogue of the result shown by Krieg [Kr 86] in the case of Siegel modular forms of integral weight.

1.5

This paper is organized as follows: in Sect. 2, the necessary notation and definitions are reviewed. In Sect. 3, the relation among Fourier-Jacobi coefficients of the Siegel-Eisenstein series and the Jacobi-Eisenstein series is derived, which is a modification of the result given by Boecherer [Bo 83] for special cases. In Sect. 4, a map from a subspace of Jacobi forms of integral weight of matrix index to a subspace of Jacobi forms of half-integral weight of integer index is defined. Moreover, the compatibility of this map with index-shift maps is studied. In Sect. 5, we calculate the action of index-shift maps on the Jacobi-Eisenstein series. We express these functions as summations of exponential functions with generalized Gauss sums. In Sect. 6, a commutativity between index-shift maps on Jacobi forms and Siegel Φ-operators is derived. In Sect. 7, a generalized Maass relation for generalized Cohen-Eisenstein series (Theorem 7.6) will be proved, while we will give a generalized Maass relation for Siegel cusp forms of half-integral weight and the proof of the main result (Theorem 8.3) in Sect. 8. We shall explain some numerical examples of the non-vanishing of the lift in Sect. 9.

Acknowledgement. The construction of the lift in this article was suggested by Professor Tamotsu Ikeda to the author at the Hakuba Autumn Workshop 2001. The author wishes to express his hearty gratitude to Professor Ikeda for the suggestion. The author also would like to express his sincere gratitude to Professor Tomoyoshi Ibukiyama for continuous encouragement. The author thanks very much to the referee, whose advice was helpful in improving the original version of the manuscript. This work was supported by JSPS KAKENHI Grant Number 23740018 and 80597766.

2 Notation and definitions

$\mathbb{R}^+$: the set of all positive real numbers
$R^{(n,m)}$: the set of $n \times m$ matrices with entries in a commutative ring R
Sym_n^* : the set of all half-integral symmetric matrices of size n
Sym_n^+ : all positive definite matrices in Sym_n^*
${}^tB$: the transpose of a matrix B
$A[B] := {}^tBAB$ for two matrices $A \in R^{(n,n)}$ and $B \in R^{(n,m)}$
1_n (resp. 0_n) : identity matrix (resp. zero matrix) of size n
$\mathrm{tr}(S)$: the trace of a square matrix S
$e(S) := e^{2\pi\sqrt{-1}\,\mathrm{tr}(S)}$ for a square matrix S
$\mathrm{rank}_p(x)$: the rank of matrix $x \in \mathbb{Z}^{(n,m)}$ over the finite field $\mathbb{Z}/p\mathbb{Z}$
$\mathrm{diag}(a_1, ..., a_n)$: the diagonal matrix $\begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix}$ for square matrices a_1, ..., a_n
$\left(\frac{*}{p}\right)$: the Legendre symbol for odd prime p
$\left(\frac{*}{2}\right) := 0, 1, -1$ accordingly as a is even, $a \equiv \pm 1 \bmod 8$ or $a \equiv \pm 3 \bmod 8$
$M_{k-\frac{1}{2}}(\Gamma_0^{(n)}(4))$: the space of Siegel modular forms of weight $k - \frac{1}{2}$ of degree n
$M^+_{k-\frac{1}{2}}(\Gamma_0^{(n)}(4))$: the plus-space of $M_{k-\frac{1}{2}}(\Gamma_0^{(n)}(4))$ (cf. [Ib 92]).
$\mathfrak{H}_n$: the Siegel upper half space of degree n
$\delta(\mathcal{S}) := 1$ or 0 accordingly as the statement $\mathcal{S}$ is true or false.

For any function F and operators T_1, T_2, ... , T_n we put

$$F|(T_1,T_2,...,T_n) \quad := \quad (F|T_1, F|T_2, ..., F|T_n).$$

2.1 Jacobi group

For a positive integer n we define the following groups:

$$\begin{aligned}
\mathrm{GSp}_n^+(\mathbb{R}) &:= \Big\{ g \in \mathbb{R}^{(2n,2n)} \,|\, g \left(\begin{smallmatrix} 0_n & -1_n \\ 1_n & 0_n \end{smallmatrix}\right) {}^t g = n(g) \left(\begin{smallmatrix} 0_n & -1_n \\ 1_n & 0_n \end{smallmatrix}\right) \\
&\qquad \text{for some } n(g) \in \mathbb{R}^+ \Big\}, \\
\mathrm{Sp}_n(\mathbb{R}) &:= \left\{ g \in \mathrm{GSp}_n^+(\mathbb{R}) \,|\, n(g) = 1 \right\}, \\
\Gamma_n &:= \mathrm{Sp}_n(\mathbb{R}) \cap \mathbb{Z}^{(2n,2n)}, \\
\Gamma_\infty^{(n)} &:= \left\{ \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \Gamma_n \,\middle|\, C = 0_n \right\}, \\
\Gamma_0^{(n)}(4) &:= \left\{ \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \Gamma_n \,\middle|\, C \equiv 0 \mod 4 \right\}.
\end{aligned}$$

For a matrix $g \in \mathrm{GSp}_n^+(\mathbb{R})$, the number $n(g)$ in the above definition of $\mathrm{GSp}_n^+(\mathbb{R})$ is called the *similitude* of the matrix g.
For positive integers n and r, we define a subgroup $G_{n,r}^J \subset \mathrm{GSp}_{n+r}^+(\mathbb{R})$ by

$$G_{n,r}^J \quad := \quad \left\{ \begin{pmatrix} A & & B & \\ & U & & \\ C & & D & \\ & & & V \end{pmatrix} \begin{pmatrix} 1_n & & & \mu \\ {}^t\lambda & 1_r & {}^t\mu & {}^t\lambda\mu + \kappa \\ & & 1_n & -\lambda \\ & & & 1_r \end{pmatrix} \in \mathrm{GSp}_{n+r}^+(\mathbb{R}) \right\},$$

where the matrices runs over $\begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{GSp}_n^+(\mathbb{R})$, $\begin{pmatrix} U & 0 \\ 0 & V \end{pmatrix} \in \mathrm{GSp}_r^+(\mathbb{R})$, $\lambda, \mu \in \mathbb{R}^{(n,r)}$ and $\kappa = {}^t\kappa \in \mathbb{R}^{(r,r)}$.
We will abbreviate such an element $\left(\begin{smallmatrix} A & & B & \\ & U & & \\ C & & D & \\ & & & V \end{smallmatrix}\right) \left(\begin{smallmatrix} 1_n & & & \mu \\ {}^t\lambda & 1_r & {}^t\mu & {}^t\lambda\mu+\kappa \\ & & 1_n & -\lambda \\ & & & 1_r \end{smallmatrix}\right)$ as

$$\left(\begin{pmatrix} A & B \\ C & D \end{pmatrix} \times \begin{pmatrix} U & 0 \\ 0 & V \end{pmatrix}, [(\lambda,\mu),\kappa] \right).$$

We remark that two matrices $\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} U & 0 \\ 0 & V \end{smallmatrix}\right)$ in the above notation have the same similitude. We will often write

$$\left(\begin{pmatrix} A & B \\ C & D \end{pmatrix}, [(\lambda,\mu),\kappa] \right)$$

instead of writing $(\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right) \times 1_{2r}, [(\lambda,\mu),\kappa])$ for simplicity. We remark that the element $(\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right), [(\lambda,\mu),\kappa])$ belongs to $\mathrm{Sp}_{n+r}(\mathbb{R})$. Similarly, an element

$$\left(\begin{smallmatrix} 1_n & & & \mu \\ {}^t\lambda & 1_r & {}^t\mu & {}^t\lambda\mu+\kappa \\ & & 1_n & -\lambda \\ & & & 1_r \end{smallmatrix}\right) \left(\begin{smallmatrix} A & & B & \\ & U & & \\ C & & D & \\ & & & V \end{smallmatrix}\right)$$

will be abbreviated as

$$\left([(\lambda,\mu),\kappa],\begin{pmatrix} A & B \\ C & D\end{pmatrix}\times\begin{pmatrix} U & 0 \\ 0 & V\end{pmatrix}\right),$$

and we will abbreviate it as $([(\lambda,\mu),\kappa],\left(\begin{smallmatrix} A & B \\ C & D\end{smallmatrix}\right))$ for the case $U = V = 1_r$.
We set a subgroup $\Gamma^J_{n,r}$ of $G^J_{n,r}$ by

$$\Gamma^J_{n,r} \quad := \quad \left\{(M,[(\lambda,\mu),\kappa]) \in G^J_{n,r} \,\middle|\, M \in \Gamma_n, \lambda,\mu \in \mathbb{Z}^{(n,r)}, \kappa \in \mathbb{Z}^{(r,r)}\right\}.$$

2.2 Groups $\widetilde{\mathrm{GSp}_n^+}(\mathbb{R})$ and $\widetilde{G^J_{n,1}}$

We denote by $\widetilde{\mathrm{GSp}_n^+}(\mathbb{R})$ the group which consists of pairs $(M,\varphi(\tau))$, where M is a matrix $M = \left(\begin{smallmatrix} A & B \\ C & D\end{smallmatrix}\right) \in \mathrm{GSp}_n^+(\mathbb{R})$, and where φ is any holomorphic function on $\mathfrak{H}_n$ such that $|\varphi(\tau)|^2 = \det(M)^{-\frac{1}{2}}|\det(C\tau + D)|$. The group operation on $\widetilde{\mathrm{GSp}_n^+}(\mathbb{R})$ is given by $(M,\varphi(\tau))(M',\varphi'(\tau)) := (MM',\varphi(M'\tau)\varphi'(\tau))$.
We embed $\Gamma_0^{(n)}(4)$ into the group $\widetilde{\mathrm{GSp}_n^+}(\mathbb{R})$ via $M \to (M,\theta^{(n)}(M\tau)\,\theta^{(n)}(\tau)^{-1})$, where $\theta^{(n)}(\tau) := \displaystyle\sum_{p\in\mathbb{Z}^{(n,1)}} e(\tau[p])$ is the theta constant. We denote by $\Gamma_0^{(n)}(4)^*$ the image of $\Gamma_0^{(n)}(4)$ in $\widetilde{\mathrm{GSp}_n^+}(\mathbb{R})$ by this embedding.
We define a Heisenberg group

$$H_{n,1}(\mathbb{R}) \quad := \quad \left\{(1_{2n},[(\lambda,\mu),\kappa]) \in \mathrm{Sp}_{n+1}(\mathbb{R}) \,|\, \lambda,\mu \in \mathbb{R}^{(n,1)}, \kappa \in \mathbb{R}\right\}.$$

If there is no confusion, we will write $[(\lambda,\mu),\kappa]$ for the element $(1_{2n},[(\lambda,\mu),\kappa])$ for simplicity.
We define a group

$$\begin{aligned}\widetilde{G^J_{n,1}} \quad &:= \quad \widetilde{\mathrm{GSp}_n^+}(\mathbb{R}) \ltimes H_{n,1}(\mathbb{R}) \\ &= \quad \left\{(\tilde{M},[(\lambda,\mu),\kappa]) \,\middle|\, \tilde{M} \in \widetilde{\mathrm{GSp}_n^+}(\mathbb{R}), [(\lambda,\mu),\kappa] \in H_{n,1}(\mathbb{R})\right\}.\end{aligned}$$

Here the group operation on $\widetilde{G^J_{n,1}}$ is given by

$$(\tilde{M}_1,[(\lambda_1,\mu_1),\kappa_1])\cdot(\tilde{M}_2,[(\lambda_2,\mu_2),\kappa_2]) \quad := \quad (\tilde{M}_1\tilde{M}_2,[(\lambda',\mu'),\kappa'])$$

for $(\tilde{M}_i,[(\lambda_i,\mu_i),\kappa_i]) \in \widetilde{G^J_{n,1}}$ $(i = 1,2)$, and where $[(\lambda',\mu'),\kappa'] \in H_{n,1}(\mathbb{R})$ is the matrix determined through the identity

$$\begin{aligned}&(M_1 \times \left(\begin{smallmatrix} n(M_1) & 0 \\ 0 & 1\end{smallmatrix}\right),[(\lambda_1,\mu_1),\kappa_1])(M_2 \times \left(\begin{smallmatrix} n(M_2) & 0 \\ 0 & 1\end{smallmatrix}\right),[(\lambda_2,\mu_2),\kappa_2]) \\ &= (M_1M_2 \times \left(\begin{smallmatrix} n(M_1)n(M_2) & 0 \\ 0 & 1\end{smallmatrix}\right),[(\lambda',\mu'),\kappa'])\end{aligned}$$

in $G^J_{n,1}$. Here $n(M_i)$ is the similitude of M_i.

2.3 Action of the Jacobi group

The group $G^J_{n,r}$ acts on $\mathfrak{H}_n \times \mathbb{C}^{(n,r)}$ by

$$\gamma\cdot(\tau,z) \quad := \quad \left(\begin{pmatrix} A & B \\ C & D\end{pmatrix}\cdot\tau\,,\ {}^t(C\tau+D)^{-1}(z+\tau\lambda+\mu){}^tU\right)$$

for any $\gamma = ((\begin{smallmatrix} A & B \\ C & D\end{smallmatrix})\times(\begin{smallmatrix} U & 0 \\ 0 & V\end{smallmatrix}),[(\lambda,\mu),\kappa]) \in G^J_{n,r}$ and for any $(\tau,z)\in\mathfrak{H}_n\times\mathbb{C}^{(n,r)}$. Here $\begin{pmatrix} A & B \\ C & D\end{pmatrix}\cdot\tau := (A\tau+B)(C\tau+D)^{-1}$ is the usual transformation.

The group $\widetilde{G^J_{n,1}}$ acts on $\mathfrak{H}_n\times\mathbb{C}^{(n,1)}$ through the projection $\widetilde{G^J_{n,1}}\to G^J_{n,1}$. It means $\widetilde{G^J_{n,1}}$ acts on $\mathfrak{H}_n\times\mathbb{C}^{(n,1)}$ by

$$\tilde\gamma\cdot(\tau,z) \quad := \quad (M\times(\begin{smallmatrix} n(M) & 0 \\ 0 & 1\end{smallmatrix}),[(\lambda,\mu),\kappa])\cdot(\tau,z)$$

for $\tilde\gamma = ((M,\varphi),[(\lambda,\mu),\kappa])\in\widetilde{G^J_{n,1}}$ and for $(\tau,z)\in\mathfrak{H}_n\times\mathbb{C}^{(n,1)}$. Here $n(M)$ is the similitude of $M\in\mathrm{GSp}^+_n(\mathbb{R})$.

2.4 Factors of automorphy

Let k be an integer and let $\mathcal{M}\in\mathrm{Sym}^+_r$. For $\gamma = ((\begin{smallmatrix} A & B \\ C & D\end{smallmatrix})\times(\begin{smallmatrix} U & 0 \\ 0 & V\end{smallmatrix}),[(\lambda,\mu),\kappa])\in G^J_{n,r}$ we define a factor of automorphy

$$\begin{aligned}
&J_{k,\mathcal{M}}\,(\gamma,(\tau,z)) \\
&:= \det(V)^k\det(C\tau+D)^k\,e(V^{-1}\mathcal{M}U(((C\tau+D)^{-1}C)[z+\tau\lambda+\mu])) \\
&\qquad \times e(-V^{-1}\mathcal{M}U({}^t\lambda\tau\lambda+{}^tz\lambda+{}^t\lambda z+{}^t\mu\lambda+{}^t\lambda\mu+\kappa)).
\end{aligned}$$

We define a slash operator $|_{k,\mathcal{M}}$ by

$$(\phi|_{k,\mathcal{M}}\gamma)(\tau,z) \quad := \quad J_{k,\mathcal{M}}(\gamma,(\tau,z))^{-1}\phi(\gamma\cdot(\tau,z))$$

for any function ϕ on $\mathfrak{H}_n\times\mathbb{C}^{(n,r)}$ and for any $\gamma\in G^J_{n,r}$. We remark that

$$\begin{aligned}
J_{k,\mathcal{M}}(\gamma_1\gamma_2,(\tau,z)) &= J_{k,\mathcal{M}}(\gamma_1,\gamma_2\cdot(\tau,z))J_{k,V_1^{-1}\mathcal{M}U_1}(\gamma_2,(\tau,z)), \\
\phi|_{k,\mathcal{M}}\gamma_1\gamma_2 &= (\phi|_{k,\mathcal{M}}\gamma_1)|_{k,V_1^{-1}\mathcal{M}U_1}\gamma_2.
\end{aligned}$$

for any $\gamma_i = \left(M_i\times(\begin{smallmatrix} U_i & 0 \\ 0 & V_i\end{smallmatrix}),[(\lambda_i,\mu_i),\kappa_i]\right)\in G^J_{n,r}$ $(i=1,2)$.
Let k and m be integers. We define a slash operator $|_{k-\frac12,m}$ for any function ϕ on $\mathfrak{H}_n\times\mathbb{C}^{(n,1)}$ by

$$\phi|_{k-\frac12,m}\tilde\gamma \quad := \quad J_{k-\frac12,m}(\tilde\gamma,(\tau,z))^{-1}\phi(\tilde\gamma\cdot(\tau,z))$$

for any $\tilde\gamma = ((M,\varphi),[(\lambda,\mu),\kappa])\in\widetilde{G^J_{n,1}}$. Here we define a factor of automorphy

$$\begin{aligned}
J_{k-\frac12,m}(\tilde\gamma,(\tau,z)) \quad := \quad & \varphi(\tau)^{2k-1}e(n(M)m(((C\tau+D)^{-1}C)[z+\tau\lambda+\mu])) \\
& \times e(-n(M)m({}^t\lambda\tau\lambda+{}^tz\lambda+{}^t\lambda z+{}^t\mu\lambda+{}^t\lambda\mu+\kappa)),
\end{aligned}$$

where $n(M)$ is the similitude of M. We remark that

$$\begin{aligned} J_{k-\frac{1}{2},m}(\tilde{\gamma_1}\tilde{\gamma_2},(\tau,z)) &= J_{k-\frac{1}{2},m}(\tilde{\gamma_1},\tilde{\gamma_2}\cdot(\tau,z))J_{k-\frac{1}{2},n(M_1)m}(\tilde{\gamma_2},(\tau,z)) \\ \phi|_{k-\frac{1}{2},m}\tilde{\gamma_1}\tilde{\gamma_2} &= (\phi|_{k-\frac{1}{2},m}\tilde{\gamma_1})|_{k-\frac{1}{2},n(M_1)m}\tilde{\gamma_2} \end{aligned}$$

for any $\tilde{\gamma_i} = ((M_i,\varphi_i),[(\lambda_i,\mu_i),\kappa_i]) \in \widetilde{G^J_{n,1}}$ $(i=1,2)$.

2.5 Jacobi forms of matrix index

We quote the definition of Jacobi form of matrix index from [Zi 89].

Definition 1. *For an integer k and for an matrix $\mathcal{M} \in Sym_r^+$, a $\mathbb{C}$-valued holomorphic function ϕ on $\mathfrak{H}_n \times \mathbb{C}^{(n,r)}$ is called a Jacobi form of weight k of index $\mathcal{M}$ of degree n, if ϕ satisfies the following two conditions:*

1. *the transformation formula $\phi|_{k,\mathcal{M}}\gamma = \phi$ for any $\gamma \in \Gamma^J_{n,r}$,*
2. *ϕ has the Fourier expansion: $\phi(\tau,z) = \sum_{\substack{N\in Sym_n^*, R\in Z^{(n,r)} \\ 4N-R\mathcal{M}^{-1\,t}R\geq 0}} c(N,R)e(N\tau)e({}^tRz)$.*

We remark that the second condition follows from the Koecher principle (cf. [Zi 89, Lemma 1.6]) if $n>1$. In the condition (2), if ϕ satisfies $c(N,R)=0$ unless $4N - R\mathcal{M}^{-1t}R > 0$, then ϕ is called a *Jacobi cusp form*.
We denote by $J^{(n)}_{k,\mathcal{M}}$ the $\mathbb{C}$-vector space of Jacobi forms of weight k of index $\mathcal{M}$ of degree n.

2.6 Jacobi forms of half-integral weight

We set a subgroup $\Gamma^{J*}_{n,1}$ of $\widetilde{G^J_{n,1}}$ by

$$\begin{aligned} \Gamma^{J*}_{n,1} &:= \left\{(M^*,[(\lambda,\mu),\kappa]) \in \widetilde{G^J_{n,1}} \,|\, M^* \in \Gamma_0^{(n)}(4)^*,\ \lambda,\mu\in\mathbb{Z}^{(n,1)}, \kappa\in\mathbb{Z}\right\} \\ &\cong \Gamma_0^{(n)}(4)^* \ltimes H_{n,1}(\mathbb{Z}), \end{aligned}$$

where we put $H_{n,1}(\mathbb{Z}) := H_{n,1}(\mathbb{R}) \cap \mathbb{Z}^{(2n+2,2n+2)}$. Here the group $\Gamma_0^{(n)}(4)^*$ was defined in §2.2.

Definition 2. *For an integer k and for an integer m, a holomorphic function ϕ on $\mathfrak{H}_n \times \mathbb{C}^{(n,1)}$ is called a Jacobi form of weight $k-\frac{1}{2}$ of index m, if ϕ satisfies the following two conditions:*

1. *the transformation formula $\phi|_{k-\frac{1}{2},m}\gamma^* = \phi$ for any $\gamma^* \in \Gamma^{J*}_{n,1}$,*
2. *$\phi^2|_{2k-1,2m}\gamma$ has the Fourier expansion for any $\gamma \in \Gamma^J_{n,1}$:*

$$\left(\phi^2|_{2k-1,2m}\gamma\right)(\tau,z) = \sum_{\substack{N\in Sym_n^*, R\in\mathbb{Z}^{(n,1)} \\ 4Nm-hR^tR\geq 0}} C(N,R)\, e\!\left(\frac{1}{h}N\tau\right) e\!\left({}^tRz\right).$$

with a integer $h > 0$, and where the slash operator $|_{k-\frac{1}{2},m}$ was defined in §2.4.

In the condition (2), for any γ if ϕ satisfies $C(N,R) = 0$ unless $4Nm - hR^tR > 0$, then ϕ is called a *Jacobi cusp form*.
We denote by $J^{(n)}_{k-\frac{1}{2},m}$ the $\mathbb{C}$-vector space of Jacobi forms of weight $k - \frac{1}{2}$ of index m of degree n.

2.7 Index-shift maps of Jacobi forms

In this subsection we introduce two kinds of maps. The both maps shift the index of Jacobi forms and these are generalizations of the V_l-map in the sense of Eichler-Zagier [E-Z 85].
We define two groups $\mathrm{GSp}_n^+(\mathbb{Z}) := \mathrm{GSp}_n^+(\mathbb{R}) \cap \mathbb{Z}^{(2n,2n)}$ and

$$\widetilde{\mathrm{GSp}_n^+(\mathbb{Z})} \quad := \quad \left\{ (M,\varphi) \in \widetilde{\mathrm{GSp}_n^+(\mathbb{R})} \,\middle|\, M \in \mathrm{GSp}_n^+(\mathbb{Z}) \right\}.$$

First we define index-shift maps for Jacobi forms of *integral weight of matrix index*. Let $\mathcal{M} = \left(\begin{smallmatrix} * & * \\ * & 1 \end{smallmatrix}\right) \in \mathrm{Sym}_2^+$. Let $X \in \mathrm{GSp}_n^+(\mathbb{Z})$ be a matrix such that the similitude of X is $n(X) = p^2$ with a prime p. For any $\phi \in J^{(n)}_{k,\mathcal{M}}$ we define the map

$$\begin{aligned} &\phi|V(X) \\ &:= \sum_{u,v \in (\mathbb{Z}/p\mathbb{Z})^{(n,1)}} \sum_{M \in \Gamma_n \backslash \Gamma_n X \Gamma_n} \phi|_{k,\mathcal{M}} \left(M \times \left(\begin{smallmatrix} p^2 & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & p \end{smallmatrix}\right), [((0,u),(0,v)),0_2] \right), \end{aligned}$$

where $(0,u),(0,v) \in (\mathbb{Z}/p\mathbb{Z})^{(n,2)}$ and where 0_2 is the zero matrix of size 2. See §2.1 for the symbol of the matrix $\left(M \times \left(\begin{smallmatrix} p^2 & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & p \end{smallmatrix}\right), [((0,u),(0,v)),0_2] \right)$.
The above summations are finite sums and do not depend on the choice of the representatives u, v and M. A straightforward calculation shows that $\phi|V(X)$ belongs to $J^{(n)}_{k,\mathcal{M}[\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)]}$. Namely $V(X)$ is a map:

$$V(X) \ : \ J^{(n)}_{k,\mathcal{M}} \to J^{(n)}_{k,\mathcal{M}[\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)]}.$$

For the sake of simplicity we set

$$V_{\alpha,n-\alpha}(p^2) \quad := \quad V(\mathrm{diag}(1_\alpha, p1_{n-\alpha}, p^2 1_\alpha, p1_{n-\alpha}))$$

for any prime p and for any α $(0 \le \alpha \le n)$.
Next we shall define index-shift maps for Jacobi forms of *half-integral weight of integer index*. We assume that p is *an odd prime*. Let m be a positive integer.

Let $Y=(X,\varphi)\in\widetilde{\mathrm{GSp}_n^+(\mathbb{Z})}$ with $n(X)=p^{2t}$, where t is a positive integer. For $\psi\in J^{(n)}_{k-\frac12,m}$ we define

$$\psi|\widetilde{V}(Y) \quad := \quad n(X)^{\frac{n(2k-1)}{4}-\frac{n(n+1)}{2}} \sum_{\tilde{M}\in\Gamma_0^{(n)}(4)^*\backslash\Gamma_0^{(n)}(4)^*Y\Gamma_0^{(n)}(4)^*} \psi|_{k-\frac12,m}(\tilde{M},[(0,0),0]),$$

where the above summation is a finite sum and does not depend on the choice of the representatives $\tilde{M}$. A direct computation shows that $\psi|\widetilde{V}(Y)$ belongs to $J^{(n)}_{k-\frac12,mp^{2t}}$.

For the sake of simplicity we set

$$\tilde{V}_{\alpha,n-\alpha}(p^2) \quad := \quad \tilde{V}((\mathrm{diag}(1_\alpha,p1_{n-\alpha},p^21_\alpha,p1_{n-\alpha}),p^{\alpha/2}))$$

for any odd prime p and for any α $(0\le\alpha\le n)$. As for $p=2$, we will introduce index-shift maps $\tilde{V}_{\alpha,n-\alpha}(4)$ in §4.7, which are maps from a subspace $J^{(n)*}_{k-\frac12,m}$ of $J^{(n)}_{k-\frac12,m}$ to $J^{(n)}_{k-\frac12,4m}$.

2.8 Hecke operators for Siegel modular forms of half-integral weight

The Hecke theory for Siegel modular forms was first introduced by Shimura [Sh 73] for degree $n=1$ and by Zhuravlev [Zh 83, Zh 84] for degree $n>1$. We quote the definition of Hecke operator from [Zh 83, Zh 84]. Let $Y=(X,\varphi)\in\widetilde{\mathrm{GSp}_n^+(\mathbb{Z})}$. Let $\phi\in M_{k-\frac12}(\Gamma_0^{(n)}(4))$. We define

$$\phi|\tilde{T}(Y) \quad := \quad n(X)^{\frac{n(2k-1)}{4}-\frac{n(n+1)}{2}} \sum_{\tilde{M}\in\Gamma_0^{(n)}(4)^*\backslash\Gamma_0^{(n)}(4)^*Y\Gamma_0^{(n)}(4)^*} \phi|_{k-\frac12}\tilde{M},$$

where $(\phi|_{k-\frac12}\tilde{M})(\tau):=\varphi(\tau)^{-2k+1}\phi(M\cdot\tau)$ for $\tilde{M}=(M,\varphi)$ and $n(X)$ is the similitude of X. For the sake of simplicity we set

$$\tilde{T}_{\alpha,n-\alpha}(p^2) \quad := \quad \tilde{T}((\mathrm{diag}(1_\alpha,p1_{n-\alpha},p^21_\alpha,p1_{n-\alpha}),p^{\alpha/2}))$$

for any odd prime p and for any α $(0\le\alpha\le n)$.

2.9 L-function of Siegel modular forms of half-integral weight

In this subsection we review the Hecke theory for Siegel modular forms of half-integral weight which has been introduced by Zhuravlev [Zh 83, Zh 84] and quote the definition of L-function of a Siegel modular form of half-integral weight.

Let $\tilde{\mathcal{H}}^{(m)}_{p^2}$ be the local Hecke ring generated by double cosets

$$K^{(m)}_\alpha := \Gamma_0^{(m)}(4)^*(\mathrm{diag}(1_\alpha,p1_{m-\alpha},p^21_\alpha,p1_{m-\alpha}),p^{\alpha/2})\Gamma_0^{(m)}(4)^* \quad (0\le\alpha\le m)$$

and ${K_0^{(m)}}^{-1}$ over $\mathbb{C}$. If p is an odd prime, then it is shown in [Zh 83, Zh 84] that the local Hecke ring $\tilde{\mathcal{H}}_{p^2}^{(m)}$ is commutative and there exists the isomorphism map

$$\Psi_m \ : \ \tilde{\mathcal{H}}_{p^2}^{(m)} \quad \to \quad R_m,$$

where the symbol R_m denotes $R_m := \mathbb{C}^{W_2}\big[z_0^{\pm}, z_1^{\pm}, ..., z_m^{\pm}\big]$, and where the subring $\mathbb{C}^{W_2}\big[z_0^{\pm}, z_1^{\pm}, ..., z_m^{\pm}\big]$ of $\mathbb{C}\left[z_0^{\pm}, z_1^{\pm}, ..., z_m^{\pm}\right]$ consists of all W_2-invariant polynomials. Here W_2 is the Weyl group of a symplectic group and the action of W_2 on $\mathbb{C}\big[z_0^{\pm}, ..., z_m^{\pm}\big]$ is generated by all permutations of $\{z_1, ..., z_m\}$ and by the maps

$$\sigma_i \ : \ z_0 \to z_0 z_i, \ z_i \to z_i^{-1}, \ z_j \to z_j \ (j \neq i)$$

for $i = 1, ..., m$. The isomorphism Ψ_m is defined as follows: Let

$$T \quad = \quad \sum_i a_i \Gamma_0^{(m)}(4)^*(X_i, \varphi_i)$$

be a decomposition of $T \in \tilde{\mathcal{H}}_{p^2}^{(m)}$, where $a_i \in \mathbb{C}$ and $(X_i, \varphi_i) \in \widetilde{\mathrm{GSp}_n^+(\mathbb{Z})}$. We can assume that X_i is an upper-triangular matrix $X_i = \begin{pmatrix} p^{\delta_i\, t}D_i{}^{-1} & B_i \\ 0 & D_i \end{pmatrix}$ with

$$D_i = \begin{pmatrix} d_{i1} & * & * \\ 0 & \ddots & * \\ 0 & 0 & d_{im} \end{pmatrix}$$

and φ_i is a constant function. It is known that $|\varphi_i|^{-1}\varphi_i$ is a forth root of unity. Then $\Psi_m(T)$ is given by

$$\Psi_m(T) \quad := \quad \sum_i a_i \left(\frac{\varphi_i}{|\varphi_i|}\right)^{-2k+1} z_0^{\delta_i} \prod_{j=1}^m (p^{-j} z_j)^{d_{ij}}$$

with a fixed integer k. For the explicit decomposition of generators $K_\alpha^{(m)}$ by left $\Gamma_0^{(m)}(4)^*$-cosets, see [Zh 83, Prop.7.1].
We define $\gamma_j \in \mathbb{C}[z_1^{\pm}, ..., z_m^{\pm}]$ $(j = 0, ..., 2m)$ through the identity

$$\sum_{j=0}^{2m} \gamma_j X^j \quad = \quad \prod_{i=1}^m \left\{ (1 - z_i X)\,(1 - z_i^{-1} X) \right\}.$$

Here γ_j $(j = 0, ..., 2m)$ is a W_2-invariant. There exists $\tilde{\gamma}_{i,p} \in \tilde{\mathcal{H}}_{p^2}^{(m)}$ $(i = 0, ..., 2m)$ which satisfies $\Psi_m(\tilde{\gamma}_{i,p}) = \gamma_i \in R_m$. We remark that $\tilde{\gamma}_{i,p} = \tilde{\gamma}_{2m-i,p}$ and $\tilde{\gamma}_{0,p} = K_0^{(m)}$.

For $p=2$ we will introduce in §4.3 the Hecke operators $\tilde{T}_{\alpha,m-\alpha}(4)$ $(\alpha=0,...,m)$ through the isomorphism between Siegel modular forms of half-integral weight and Jacobi forms of index 1 (see (4.2) in §4.3). We remark that the Hecke operators $\tilde{T}_{\alpha,m-\alpha}(4)$ $(\alpha=0,...,m)$ are defined for the generalized plus-space, which is a subspace of $M_{k-\frac{1}{2}}(\Gamma_0^{(m)}(4))$. Through the definition of $\tilde{\gamma}_{i,p}$ for odd prime p, we define $\tilde{\gamma}_{i,2}$ in the same formula by using $\tilde{T}_{\alpha,m-\alpha}(4)$ $(\alpha=0,...,m)$ as in the case of odd primes. by replacing p by 2.
Let $F \in M^+_{k-\frac{1}{2}}(\Gamma_0^{(m)}(4))$ be an eigenform for any Hecke operator $\tilde{T}_{\alpha,m-\alpha}(p^2)$ $(0 \le \alpha \le m)$ and for any prime p. Here $M^+_{k-\frac{1}{2}}(\Gamma_0^{(m)}(4))$ denotes the generalized plus-space which is a subspace of $M_{k-\frac{1}{2}}(\Gamma_0^{(m)}(4))$ (see [Ib 92] or §4.3 for the definition of $M^+_{k-\frac{1}{2}}(\Gamma_0^{(m)}(4))$). We define the Euler p-factor of F by

$$Q_{F,p}(z): \quad = \quad \sum_{j=0}^{2m} \lambda_F(\tilde{\gamma}_{j,p}) z^j,$$

where $\lambda_F(\tilde{\gamma}_{j,p})$ is the eigenvalue of F with respect to $\tilde{\gamma}_{j,p}$. There exists a set of complex numbers $\{\mu_{0,p}^2, \mu_{1,p}^{\pm}, ...\mu_{m,p}^{\pm}\}$ which satisfies

$$Q_{F,p}(z) \quad = \quad \prod_{i=1}^{m} \left\{ \left(1-\mu_{i,p}z\right)\left(1-\mu_{i,p}^{-1}z\right)\right\}$$

and

$$\mu_{0,p}^2 \mu_{1,p} \cdots \mu_{m,p} \quad = \quad p^{m(2k-1)/2 - m(m+1)/2},$$

since $\gamma_{2m-j} = \gamma_j$ $(j=0,...,m-1)$, $Q_{F,p}(z^{-1}) = z^{-2m} Q_{F,p}(z)$ and $Q_{F,p}(0) = 1 \neq 0$. Following Zhuravlev [Zh 84] we call the set $\{\mu_{0,p}^2, \mu_{1,p}^{\pm}, ..., \mu_{m,p}^{\pm}\}$ *the p-parameters* of F. The L-function of F is defined by

$$L(s,F) \quad := \quad \prod_p Q_{F,p}(p^{-s+k-3/2})^{-1}.$$

3 Fourier-Jacobi expansion of Siegel-Eisenstein series with matrix index

In this section we assume that k is an even integer.
Let r be a non-negative integer. For $\mathcal{M} \in \mathrm{Sym}_r^+$ and for an even integer k we define the Jacobi-Eisenstein series of weight k of index $\mathcal{M}$ of degree n by

$$E_{k,\mathcal{M}}^{(n)} \quad := \quad \sum_{M \in \Gamma_\infty^{(n)} \backslash \Gamma_n} \sum_{\lambda \in \mathbb{Z}^{(n,r)}} 1|_{k,\mathcal{M}}([(\lambda,0),0_r], M). \qquad (3.1)$$

The above sum converges for $k > n + r + 1$ (cf. [Zi 89]). The Siegel-Eisenstein series $E_k^{(n)}$ of weight k of degree n is defined by

$$E_k^{(n)}(Z) \quad := \sum_{\left(\begin{smallmatrix} * & * \\ C & D \end{smallmatrix}\right) \in \Gamma_\infty^{(n)} \backslash \Gamma_n} \det(CZ + D)^{-k}, \tag{3.2}$$

where $Z \in \mathfrak{H}_n$. We denote by $e_{k,\mathcal{M}}^{(n-r)}$ the $\mathcal{M}$-th Fourier-Jacobi coefficient of $E_k^{(n)}$, it means that

$$E_k^{(n)}(\left(\begin{smallmatrix} \tau & z \\ {}^t z & \omega \end{smallmatrix}\right)) \quad = \sum_{\mathcal{M} \in Sym_r^*} e_{k,\mathcal{M}}^{(n-r)}(\tau, z)\, e(\mathcal{M}\omega) \tag{3.3}$$

is a Fourier-Jacobi expansion of the Siegel-Eisenstein series $E_k^{(n)}$ of weight k of degree n, where $\tau \in \mathfrak{H}_{n-r}$, $\omega \in \mathfrak{H}_r$ and $z \in \mathbb{C}^{(n-r,r)}$. The explicit formula for the Fourier-Jacobi expansion of Siegel-Eisenstein series is given in [Bo 83, Satz 7] for arbitrary degree.

The purpose of this section is to express the Fourier-Jacobi coefficient $e_{k,\mathcal{M}}^{(n-2)}$ for $\mathcal{M} = \left(\begin{smallmatrix} * & * \\ * & 1 \end{smallmatrix}\right) \in \mathrm{Sym}_2^+$ as a summation of Jacobi-Eisenstein series of matrix index (Proposition 3.3).

We first obtain the following lemma.

Lemma 3.1. *For any $\mathcal{M} \in Sym_r^+$ and for any $A \in GL_r(\mathbb{Z})$ we have*

$$E_{k,\mathcal{M}}^{(n)}(\tau, z) \quad = \quad E_{k,\mathcal{M}[A^{-1}]}^{(n)}(\tau, z\,{}^t A)$$

and

$$e_{k,\mathcal{M}}^{(n)}(\tau, z) \quad = \quad e_{k,\mathcal{M}[A^{-1}]}^{(n)}(\tau, z\,{}^t A).$$

Proof. The first identity follows directly from the definition. The transformation formula $E_k^{(n+r)}\left(\begin{pmatrix} 1_n & \\ & A \end{pmatrix}\begin{pmatrix} \tau & z \\ {}^t z & \omega \end{pmatrix}\begin{pmatrix} 1_n & \\ & {}^t A \end{pmatrix}\right) = E_k^{(n+r)}\left(\begin{pmatrix} \tau & z \\ {}^t z & \omega \end{pmatrix}\right)$ gives the second identity. □

Let m be a positive integer. We denote by D_0 the discriminant of $\mathbb{Q}(\sqrt{-m})$, and we put $f := \sqrt{\frac{m}{|D_0|}}$. We note that f is a positive integer if $-m \equiv 0, 1$ mod 4.

We denote by $h_{k-\frac{1}{2}}(m)$ the m-th Fourier coefficient of the Cohen-Eisenstein series of weight $k - \frac{1}{2}$ (cf. Cohen [Co 75]). The following formula is known (cf. [Co 75], [E-Z 85]):

$$h_{k-\frac{1}{2}}(m) = \begin{cases} h_{k-\frac{1}{2}}(|D_0|)\, m^{k-\frac{3}{2}} \sum_{d|f} \mu(d) \left(\frac{D_0}{d}\right) d^{1-k} \sigma_{3-2k}\left(\frac{f}{d}\right) & \text{if } -m \equiv 0, 1 \bmod 4, \\ 0 & \text{otherwise,} \end{cases}$$

where we define $\sigma_a(b) := \sum_{d|b} d^a$ and μ is the Möbis function.

We assume $-m \equiv 0, 1 \bmod 4$. Let D_0 and f be as above. For the sake of simplicity we define

$$g_k(m) := \sum_{d|f} \mu(d)\, h_{k-\frac{1}{2}}\left(\frac{m}{d^2}\right).$$

We will use the following lemma for the proof of Proposition 7.5.

Lemma 3.2. *Let m be a natural number such that $-m \equiv 0,\ 1 \mod 4$. Then for any prime p we have*

$$g_k(p^2 m) \quad = \quad \left(p^{2k-3} - \left(\frac{-m}{p}\right) p^{k-2}\right) g_k(m).$$

Proof. Let D_0, f be as above. By using the formula of $h_{k-\frac{1}{2}}(m)$ we obtain

$$h_{k-\frac{1}{2}}(m) = h_{k-\frac{1}{2}}(|D_0|)|D_0|^{k-\frac{3}{2}} \prod_{q|f} \left\{ \sigma_{2k-3}(q^{l_q}) - \left(\frac{D_0}{q}\right) q^{k-2} \sigma_{2k-3}(q^{l_q-1}) \right\},$$

where q runs over all primes which divide f, and where we put $l_q := \mathrm{ord}_q(f)$. In particular, the function $h_{k-\frac{1}{2}}(m)(h_{k-\frac{1}{2}}(|D_0|)|D_0|^{k-\frac{3}{2}})^{-1}$ is multiplicative as function of f. Hence, for any prime q, we have

$$\begin{aligned}
&h_{k-\frac{1}{2}}(|D_0|q^{2l_q}) - h_{k-\frac{1}{2}}(|D_0|q^{2l_q-2}) \\
&= h_{k-\frac{1}{2}}(|D_0|)|D_0|^{k-\frac{3}{2}} \left(q^{(2k-3)l_q} - \left(\frac{D_0}{q}\right) q^{k-2+(2k-3)(l_q-1)} \right),
\end{aligned}$$

Thus

$$\begin{aligned}
g_k(m) &= h_{k-\frac{1}{2}}(|D_0|)|D_0|^{k-\frac{3}{2}} \sum_{d|f} \mu(d) \frac{h_{k-\frac{1}{2}}\left(\frac{m}{d^2}\right)}{h_{k-\frac{1}{2}}(|D_0|)|D_0|^{k-\frac{3}{2}}} \\
&= h_{k-\frac{1}{2}}(|D_0|)|D_0|^{k-\frac{3}{2}} \prod_{q|f} \frac{h_{k-\frac{1}{2}}(|D_0|q^{2l_q}) - h_{k-\frac{1}{2}}(|D_0|q^{2l_q-2})}{h_{k-\frac{1}{2}}(|D_0|)|D_0|^{k-\frac{3}{2}}} \\
&= h_{k-\frac{1}{2}}(|D_0|)|D_0|^{k-\frac{3}{2}} \prod_{q|f} \left(q^{(2k-3)l_q} - \left(\frac{D_0}{q}\right) q^{k-2+(2k-3)(l_q-1)} \right).
\end{aligned}$$

The lemma follows from this identity, since $\left(\frac{-m}{p}\right) = 0$ if $p|f$; $\left(\frac{-m}{p}\right) = \left(\frac{D_0}{p}\right)$ if $p \nmid f$. ∎

By using the function $g_k(m)$, we obtain the following proposition.

Proposition 3.3. *For* $\mathcal{M} = \begin{pmatrix} * & * \\ * & 1 \end{pmatrix} \in Sym_2^+$ *we put* $m = \det(2\mathcal{M})$. *Let* D_0, f *be as above, which depend on the integer* m. *If* $k > n+1$, *then*

$$e_{k,\mathcal{M}}^{(n-2)}(\tau, z) \quad = \quad \sum_{d|f} g_k\left(\frac{m}{d^2}\right) E_{k,\mathcal{M}[W_d{}^{-1}]}^{(n-2)}(\tau, z\,{}^tW_d),$$

where we chose a matrix $W_d \in GL_2(\mathbb{Q}) \cap \mathbb{Z}^{(2,2)}$ *for each* d *which satisfies the conditions* $\det(W_d) = d$, ${}^tW_d{}^{-1}\mathcal{M}W_d{}^{-1} \in Sym_2^+$ *and* ${}^tW_d{}^{-1}\mathcal{M}W_d{}^{-1} = \begin{pmatrix} * & * \\ * & 1 \end{pmatrix}$. *Remark that the matrix* W_d *is not uniquely determined, but the above summation does not depend on the choice of* W_d.

Proof. We use the terminology and the Satz 7 in [Bo 83] for this proof. For $\mathcal{M}' \in \mathrm{Sym}_n^+$ we denote by $a_2^k(\mathcal{M}')$ the $\mathcal{M}'$-th Fourier coefficient of Siegel-Eisenstein series of weight k of degree 2. We put

$$\mathrm{M}_2^n(\mathbb{Z})^* := \left\{ N \in \mathbb{Z}^{(2,2)} \mid \det(N) \neq 0 \text{ and there exists } V = \left(\begin{smallmatrix} N & * \\ * & * \end{smallmatrix}\right) \in \mathrm{GL}_n(\mathbb{Z}) \right\}.$$

A matrix $N \in \mathbb{Z}^{(n,2)}$ is called *primitive* if there exists a matrix $V \in \mathrm{GL}_n(\mathbb{Z})$ such that $V = (N\ *)$. From [Bo 83, Satz 7] we have

$$e_{k,\mathcal{M}}^{(n-2)}(\tau, z) \quad = \sum_{\substack{N_1 \in M_2^n(\mathbb{Z})^*/GL_2(\mathbb{Z}) \\ N_1^{-1}\mathcal{M}\,{}^tN_1{}^{-1} \in Sym_2^+}} a_2^k(\mathcal{M}[{}^tN_1{}^{-1}]) \sum_{\substack{N_3 \in \mathbb{Z}^{(n-2,2)} \\ \left(\begin{smallmatrix} N_1 \\ N_3 \end{smallmatrix}\right):primitive}} f(\mathcal{M}, N_1, N_3; \tau, z),$$

where

$$\begin{aligned} & f(\mathcal{M}, N_1, N_3; \tau, z) \\ & \quad = \sum_{\left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right) \in \Gamma_\infty^{(n-2)} \backslash \Gamma_{n-2}} \det(C\tau + D)^{-k} \\ & \qquad\qquad \times e\big(\mathcal{M}\left\{-{}^tz(C\tau+D)^{-1}Cz + {}^tz(C\tau+D)^{-1}N_3N_1^{-1}\right. \\ & \qquad\qquad + {}^tN_1{}^{-1}\,{}^tN_3\,{}^t(C\tau+D)^{-1}z \\ & \qquad\qquad \left. + {}^tN_1{}^{-1}\,{}^tN_3(A\tau+B)(C\tau+D)^{-1}N_3N_1^{-1}\right\}\big). \end{aligned}$$

For any positive integer l such that $l^2|m$, we chose a matrix $W_l \in \mathbb{Z}^{(2,2)}$ which satisfies three conditions $\det(W_l) = l$, ${}^tW_l{}^{-1}\mathcal{M}W_l{}^{-1} \in \mathrm{Sym}_2^+$ and ${}^tW_l{}^{-1}\mathcal{M}W_l{}^{-1} = \begin{pmatrix} * & * \\ * & 1 \end{pmatrix}$. By virtue of these conditions, W_l has the form $W_l = \begin{pmatrix} l & 0 \\ x & 1 \end{pmatrix}$ with some $x \in \mathbb{Z}$. The set ${}^tW_l\,\mathrm{GL}_2(\mathbb{Z})$ is uniquely determined for each positive integer l such that $l^2|m$. If $N_1 = {}^tW_l = \begin{pmatrix} l & x \\ 0 & 1 \end{pmatrix}$, then

$$\sum_{\substack{N_3 \in \mathbb{Z}^{(n-2,2)} \\ \binom{N_1}{N_3}:primitive}} f(\mathcal{M}, N_1, N_3; \tau, z) \quad = \quad \sum_{a|l} \mu(a) \sum_{N_3 \in \mathbb{Z}^{(n-2,2)}} f(\mathcal{M}, N_1, N_3 \left(\begin{smallmatrix} a & 0 \\ 0 & 1 \end{smallmatrix}\right); \tau, z).$$

Thus

$$\begin{aligned}
& e_{k,\mathcal{M}}^{(n-2)}(\tau, z) \\
&= \sum_{\substack{l \\ l^2|m}} a_2^k(\mathcal{M}[W_l{}^{-1}]) \sum_{a|l} \mu(a) \sum_{N_3 \in \mathbb{Z}^{(n-2,2)}} f(\mathcal{M}, {}^tW_l, N_3 \left(\begin{smallmatrix} a & 0 \\ 0 & 1 \end{smallmatrix}\right); \tau, z) \\
&= \sum_{\substack{l \\ l^2|m}} a_2^k(\mathcal{M}[W_l{}^{-1}]) \sum_{a|l} \mu(a) \\
&\quad \times \sum_{N_3 \in \mathbb{Z}^{(n-2,2)}} f(\mathcal{M}[W_l{}^{-1} \left(\begin{smallmatrix} a & 0 \\ 0 & 1 \end{smallmatrix}\right)], 1_2, N_3; \tau, z\,{}^tW_l \left(\begin{smallmatrix} a & 0 \\ 0 & 1 \end{smallmatrix}\right)^{-1}).
\end{aligned}$$

Therefore

$$\begin{aligned}
& e_{k,\mathcal{M}}^{(n-2)}(\tau, z) \\
&= \sum_{\substack{l \\ l^2|m}} a_2^k(\mathcal{M}[W_l{}^{-1}]) \sum_{a|l} \mu(a)\, E_{k,\mathcal{M}[W_l^{-1}\left(\begin{smallmatrix} a & \\ & 1 \end{smallmatrix}\right)]}^{(n-2)}(\tau, z\,{}^tW_l \left(\begin{smallmatrix} a^{-1} & \\ & 1 \end{smallmatrix}\right)) \\
&= \sum_{\substack{d \\ d^2|m}} E_{k,\mathcal{M}[W_d^{-1}]}^{(n-2)}(\tau, z\,{}^tW_d) \sum_{\substack{a \\ a^2|\frac{m}{d^2}}} \mu(a)\, a_2^k(\mathcal{M}[W_d{}^{-1} \left(\begin{smallmatrix} a^{-1} & \\ & 1 \end{smallmatrix}\right)]).
\end{aligned}$$

Here we have $a_2^k(\mathcal{M}') = h_{k-\frac{1}{2}}(\det(2\mathcal{M}'))$ for any $\mathcal{M}' = \left(\begin{smallmatrix} * & * \\ * & 1 \end{smallmatrix}\right) \in \mathrm{Sym}_2^+$. Moreover, if $m \not\equiv 0, 3 \mod 4$, then $h_{k-\frac{1}{2}}(m) = 0$. Hence

$$e_{k,\mathcal{M}}^{(n-2)}(\tau, z) \quad = \quad \sum_{\substack{d \\ d|f}} E_{k,\mathcal{M}[W_d^{-1}]}^{(n-2)}(\tau, z\,{}^tW_d) \sum_{\substack{a \\ a|\frac{f}{d}}} \mu(a)\, h_{k-\frac{1}{2}}\Big(\frac{m}{a^2 d^2}\Big).$$

Therefore this proposition follows. □

4 Relation between Jacobi forms of half-integral weight of integer index and Jacobi forms of integral weight of matrix index

In this section we fix a positive definite half-integral symmetric matrix $\mathcal{M} \in \mathrm{Sym}_2^+$, and we assume that $\mathcal{M}$ has the form $\mathcal{M} = \begin{pmatrix} l & \frac{1}{2}r \\ \frac{1}{2}r & 1 \end{pmatrix}$ with integers l and r.

The purpose of this section is to give a map $\iota_{\mathcal{M}}$ which is a linear map from a subspace of holomorphic functions on $\mathfrak{H}_n \times \mathbb{C}^{(n,2)}$ to a subspace of holomorphic functions on $\mathfrak{H}_n \times \mathbb{C}^{(n,1)}$. A restriction of $\iota_{\mathcal{M}}$ gives a map from a subspace $J_{k,\mathcal{M}}^{(n)*}$ of $J_{k,\mathcal{M}}^{(n)}$ to a subspace $J_{k-\frac{1}{2},\det(2\mathcal{M})}^{(n)*}$ of $J_{k-\frac{1}{2},\det(2\mathcal{M})}^{(n)}$ (cf. Lemma 4.2). Moreover, we shall show a compatibility between the map $\iota_{\mathcal{M}}$ and index-shift maps (cf. Proposition 4.3 and Proposition 4.4). Furthermore, we define index-shift maps $\tilde{V}_{\alpha,n-\alpha}(p^2)$ for $J_{k-\frac{1}{2},\det(2\mathcal{M})}^{(n)*}$ at $p = 2$ through the map $\iota_{\mathcal{M}}$ (cf. §4.7).

By virtue of the map $\iota_{\mathcal{M}}$ and by the results in this section, we can translate some relations among Jacobi forms of *half-integral weight of integer index* to relations among Jacobi forms of *integral weight of matrix index*.

4.1 An expansion of Jacobi forms of integer index

In this subsection we consider an expansion of Jacobi forms of integer index and shall introduce a subspace $J_{k,\mathcal{M}}^{(n)*} \subset J_{k,\mathcal{M}}^{(n)}$.

The symbol $J_{k,1}^{(n+1)}$ denotes the space of Jacobi forms of weight k of index 1 of degree $n+1$ (cf. §2.5).

Let $\phi_1(\tau, z) \in J_{k,1}^{(n+1)}$ be a Jacobi form. We regard $\phi_1(\tau, z)\, e(\omega)$ as a holomorphic function on $\mathfrak{H}_{n+2}$, where $\tau \in \mathfrak{H}_{n+1}$, $z \in \mathbb{C}^{(n+1,1)}$ and $\omega \in \mathfrak{H}_1$ such that $\left(\begin{smallmatrix} \tau & z \\ {}^t z & \omega \end{smallmatrix}\right) \in \mathfrak{H}_{n+2}$. We have an expansion

$$\phi_1(\tau, z)e(w) \quad = \sum_{\substack{S \in Sym_2^+ \\ S = \left(\begin{smallmatrix} * & * \\ * & 1 \end{smallmatrix}\right)}} \phi_{\mathcal{S}}(\tau', z')e(\mathcal{S}\omega'),$$

where $\tau' \in \mathfrak{H}_n$, $z' \in \mathbb{C}^{(n,2)}$ and $\omega' \in \mathfrak{H}_2$ which satisfy $\left(\begin{smallmatrix} \tau & z \\ {}^t z & \omega \end{smallmatrix}\right) = \begin{pmatrix} \tau' & z' \\ {}^t z' & \omega' \end{pmatrix} \in \mathfrak{H}_{n+2}$. Because the group $\Gamma_{n,2}^J$ (cf. § 2.1) is a subgroup of $\Gamma_{n+1,1}^J$, the form $\phi_{\mathcal{S}}$ belongs to $J_{k,\mathcal{S}}^{(n)}$. We denote this map by $\mathrm{FJ}_{1,\mathcal{S}}$, it means that we have a map

$$\mathrm{FJ}_{1,\mathcal{S}} : J_{k,1}^{(n+1)} \to J_{k,\mathcal{S}}^{(n)}.$$

By an abuse of language, we call the map $\mathrm{FJ}_{1,\mathcal{S}}$ *the Fourier-Jacobi expansion with respect to* S.

The $\mathbb{C}$-vector subspace $J_{k,\mathcal{M}}^{(n)*}$ of $J_{k,\mathcal{M}}^{(n)}$ denotes the image of $J_{k,1}^{(n+1)}$ by $\mathrm{FJ}_{1,\mathcal{M}}$, where $\mathcal{M}$ is a half-integral symmetric matrix of size 2.

4.2 Fourier-Jacobi expansion of Siegel modular forms of half-integral weight

The purpose of this subsection is to show the following lemma.

Lemma 4.1. *Let* $F\left(\left(\begin{smallmatrix}\tau & z\\ {}^tz & \omega\end{smallmatrix}\right)\right) = \sum_{m\in\mathbb{Z}} \phi_m(\tau,z)e(m\omega)$ *be a Fourier-Jacobi expansion of* $F \in M_{k-\frac{1}{2}}(\Gamma_0^{(n+1)}(4))$, *where* $\tau \in \mathfrak{H}_n$, $\omega \in \mathfrak{H}_1$ *and* $z \in \mathbb{C}^{(n,1)}$. *Then* $\phi_m \in J^{(n)}_{k-\frac{1}{2},m}$ *for any natural number* m.

Proof. Due to the definition of $J^{(n)}_{k-\frac{1}{2},m}$, we only need to show the identity

$$\theta^{(n+1)}(\gamma\cdot\left(\begin{smallmatrix}\tau & z\\ {}^tz & \omega\end{smallmatrix}\right))\,\theta^{(n+1)}(\left(\begin{smallmatrix}\tau & z\\ {}^tz & \omega\end{smallmatrix}\right))^{-1} \quad = \quad \theta^{(n)}(\left(\begin{smallmatrix}A & B\\ C & D\end{smallmatrix}\right)\cdot\tau)\,\theta^{(n)}(\tau)^{-1}$$

for any $\gamma = (\left(\begin{smallmatrix}A & B\\ C & D\end{smallmatrix}\right), [(\lambda,\mu),\kappa]) \in \Gamma^J_{n,1}$ and for any $\left(\begin{smallmatrix}\tau & z\\ {}^tz & \omega\end{smallmatrix}\right) \in \mathfrak{H}_{n+1}$ such that $\tau \in \mathfrak{H}_n$, $\omega \in \mathfrak{H}_1$. Here $\theta^{(n+1)}$ and $\theta^{(n)}$ are the theta constants (cf. §2.2).
For any $M = \left(\begin{smallmatrix}A' & B'\\ C' & D'\end{smallmatrix}\right) \in \Gamma_0^{(n+1)}(4)$ it is known that

$$\left(\theta^{(n+1)}(M\cdot Z)\,\theta^{(n+1)}(Z)^{-1}\right)^2 \quad = \quad \det(C'Z+D')\left(\frac{-4}{\det D'}\right),$$

where $Z \in \mathfrak{H}_{n+1}$. Here $\left(\frac{-4}{\det D'}\right)$ is the quadratic symbol and it is known the identity $\left(\frac{-4}{\det D'}\right) = (-1)^{\frac{\det D'-1}{2}}$. Hence, for any $\gamma = (\left(\begin{smallmatrix}A & B\\ C & D\end{smallmatrix}\right), [(\lambda,\mu),\kappa]) \in \Gamma^J_{n,1}$, we obtain

$$\left(\theta^{(n+1)}(\gamma\cdot Z)\,\theta^{(n+1)}(Z)^{-1}\right)^2 \quad = \quad \det(C\tau+D)\left(\frac{-4}{\det D}\right),$$

where $Z = \left(\begin{smallmatrix}\tau & z\\ {}^tz & \omega\end{smallmatrix}\right) \in \mathfrak{H}_{n+1}$ with $\tau \in \mathfrak{H}_n$. In particular, the holomorphic function $\frac{\theta^{(n+1)}(\gamma\cdot Z)}{\theta^{(n+1)}(Z)}$ does not depend on the choice of $z \in \mathbb{C}^{(n,1)}$ and of $\omega \in \mathfrak{H}_1$. We substitute $z = 0$ into $\frac{\theta^{(n+1)}(\gamma\cdot Z)}{\theta^{(n+1)}(Z)}$ and a straightforward calculation gives

$$\frac{\theta^{(n+1)}(\gamma\cdot\left(\begin{smallmatrix}\tau & 0\\ 0 & \omega\end{smallmatrix}\right))}{\theta^{(n+1)}(\left(\begin{smallmatrix}\tau & 0\\ 0 & \omega\end{smallmatrix}\right))} \quad = \quad \frac{\theta^{(n)}(\left(\begin{smallmatrix}A & B\\ C & D\end{smallmatrix}\right)\cdot\tau)}{\theta^{(n)}(\tau)}.$$

Hence we conclude this lemma. ☐

4.3 The map σ and the Hecke operator $\tilde{T}_{\alpha,n-\alpha}(p^2)$

In this subsection we review the isomorphism between the space of Jacobi forms of index 1 and a subspace of Siegel modular forms of half-integral weight, which has been shown by Eichler-Zagier[E-Z 85] for degree one and by Ibukiyama[Ib 92] for general degree.
Let $M^+_{k-\frac{1}{2}}(\Gamma_0^{(n)}(4))$ be the generalized plus-space introduced in [Ib 92, page 112], which is a generalization of the Kohnen plus-space for higher degrees:

$$M^+_{k-\frac{1}{2}}(\Gamma_0^{(n)}(4)) := \left\{ F \in M_{k-\frac{1}{2}}(\Gamma_0^{(n)}(4)) \;\middle|\; \begin{array}{c}\text{the coefficients } A(N) = 0 \text{ unless}\\ N + (-1)^k R\,{}^tR \in 4\,\mathrm{Sym}^*_n\\ \text{for some } R \in \mathbb{Z}^{(n,1)}\end{array} \right\}.$$

A form $F \in M_{k-\frac{1}{2}}(\Gamma_0^{(n)}(4))$ is called a Siegel cusp form if F^2 is a Siegel cusp form of weight $2k-1$. We denote by $S^+_{k-\frac{1}{2}}(\Gamma_0^{(n)}(4))$ the space of all Siegel cusp forms in $M^+_{k-\frac{1}{2}}(\Gamma_0^{(n)}(4))$.

For any even integer k, the isomorphism between $J_{k,1}^{(n)}$ (the space of Jacobi forms of weight k of index 1 of degree n) and $M^+_{k-\frac{1}{2}}(\Gamma_0^{(n)}(4))$ is shown in [E-Z 85, Theorem 5.4] for $n=1$ and in [Ib 92, Theorem 1] for $n>1$. We call this isomorphism *the Eichler-Zagier-Ibukiyama correspondence* and denote this linear map by σ which is a bijection from $J_{k,1}^{(n)}$ to $M^+_{k-\frac{1}{2}}(\Gamma_0^{(n)}(4))$ as modules over the ring of Hecke operators. By the map σ the space $S^+_{k-\frac{1}{2}}(\Gamma_0^{(n)}(4))$ is isomorphic to the space of Jacobi cusp forms $J_{k,1}^{(n)\,cusp}$. The map

$$\sigma : J_{k,1}^{(n)} \to M^+_{k-\frac{1}{2}}(\Gamma_0^{(n)}(4))$$

is given as follows: if

$$\phi(\tau,z) = \sum_{\substack{N\in Sym_n^*,\, R\in\mathbb{Z}^{(n,1)} \\ 4N-R^tR\geq 0}} C(N,R)\, e(N\tau + R^t z)$$

is a Jacobi form which belongs to $J_{k,1}^{(n)}$, then $\sigma(\phi) \in M^+_{k-\frac{1}{2}}(\Gamma_0^{(n)}(4))$ is defined by

$$\sigma(\phi)(\tau) \quad := \quad \sum_{\substack{R \bmod (2\mathbb{Z})^{(n,1)} \\ R\in\mathbb{Z}^{(n,1)}}} \ \sum_{\substack{N\in Sym_n^* \\ 4N-R^tR\geq 0}} C(N,R)\, e((4N-R^tR)\tau).$$

For the double coset $\Gamma_n \mathrm{diag}(1_\alpha, p1_{n-\alpha}, p^2 1_\alpha, p1_{n-\alpha})\Gamma_n$ and for $\phi \in J_{k,1}^{(n)}$, the Hecke operator $T^J_{\alpha,n-\alpha}(p^2)$ is defined by

$$\phi|T^J_{\alpha,n-\alpha}(p^2) \quad := \quad \sum_{\lambda,\mu\in(\mathbb{Z}/p\mathbb{Z})^n} \sum_M \phi|_{k,1}\left(M \times \left(\begin{smallmatrix} p & 0 \\ 0 & p \end{smallmatrix}\right), [(\lambda,\mu),0]\right).$$

Here, in the second summation of the RHS, the matrix M runs over all representatives of $\Gamma_n\backslash\Gamma_n \,\mathrm{diag}(1_\alpha, p1_{n-\alpha}, p^2 1_\alpha, p1_{n-\alpha})\Gamma_n$. Let $\tilde{T}_{\alpha,n-\alpha}(p^2)$ be the Hecke operator introduced in §2.8 for odd prime p, which acts on the space $M_{k-\frac{1}{2}}(\Gamma_0^{(n)}(4))$. For any odd prime p the identity

$$\sigma(\phi)|\tilde{T}_{\alpha,n-\alpha}(p^2) \quad = \quad p^{\alpha/2+k(2n+1)-(2n+7)n/2}\sigma(\phi|T^J_{\alpha,n-\alpha}(p^2)). \tag{4.1}$$

has been obtained in [Ib 92].

Through the identity (4.1) the Hecke operator $\tilde{T}_{\alpha,n-\alpha}(4)$ for $M^+_{k-\frac{1}{2}}(\Gamma_0^{(n)}(4))$ is defined. It means that we define

$$\sigma(\phi)|\tilde{T}_{\alpha,n-\alpha}(4) \quad := \quad 2^{\alpha/2+k(2n+1)-(2n+7)n/2}\sigma(\phi|T^J_{\alpha,n-\alpha}(4)). \tag{4.2}$$

4.4 A generalization of Cohen-Eisenstein series and the subspace $J^{(n)*}_{k-1/2}$

In this subsection we will introduce a subspace $J^{(n)*}_{k-\frac{1}{2},m} \subset J^{(n)}_{k-\frac{1}{2},m}$ for any integer n. Moreover, we will introduce a generalized Cohen-Eisenstein series $\mathcal{H}^{(n+1)}_{k-\frac{1}{2}}$ and will consider the Fourier-Jacobi expansion of $\mathcal{H}^{(n+1)}_{k-\frac{1}{2}}$ for any integer n.
Let $e^{(n+1)}_{k,1}$ be the first Fourier-Jacobi coefficient of Siegel-Eisenstein series $E^{(n+2)}_k$ (see (3.3) in §3 for the definition of $e^{(n+1)}_{k,1}$). It is known that $e^{(n+1)}_{k,1}$ coincides with the Jacobi-Eisenstein series $E^{(n+1)}_{k,1}$ of weight k of index 1 of degree $n+1$ (cf. [Bo 83, Satz 7]. See (3.1) in §3 for the definition of $E^{(n+1)}_{k,1}$).
We define the *generalized Cohen-Eisenstein series* $\mathcal{H}^{(n+1)}_{k-\frac{1}{2}}$ of weight $k-\frac{1}{2}$ of degree $n+1$ by

$$\mathcal{H}^{(n+1)}_{k-\frac{1}{2}} \quad := \quad \sigma(E^{(n+1)}_{k,1}).$$

Because $E^{(n+1)}_{k,1} \in J^{(n+1)}_{k,1}$, we have $\mathcal{H}^{(n+1)}_{k-\frac{1}{2}} \in M^+_{k-\frac{1}{2}}(\Gamma^{(n+1)}_0(4))$ for any integer n.
For any integer m we denote by $\widetilde{\mathrm{FJ}}_m$ the linear map from $M_{k-\frac{1}{2}}(\Gamma^{(n+1)}_0(4))$ to $J^{(n)}_{k-\frac{1}{2},m}$ obtained by the Fourier-Jacobi expansion with respect to the index m. It means that if $G \in M_{k-\frac{1}{2}}(\Gamma^{(n+1)}_0(4))$, then G has the expansion

$$G\left(\begin{pmatrix} \tau & z \\ {}^t z & \omega \end{pmatrix}\right) \quad = \quad \sum_{m\in\mathbb{Z}} \phi_m(\tau,z)e(m\omega)$$

and we define $\widetilde{\mathrm{FJ}}_m(G) := \phi_m$. We remark $\phi_m \in J^{(n)}_{k-\frac{1}{2},m}$ due to Lemma 4.1.
We denote by $J^{(n)*}_{k-\frac{1}{2},m}$ the image of $M^+_{k-\frac{1}{2}}(\Gamma^{(n+1)}_0(4))$ by the map $\widetilde{\mathrm{FJ}}_m$.
We denote by $e^{(n)}_{k-\frac{1}{2},m}$ the m-th Fourier-Jacobi coefficient of the generalized Cohen-Eisenstein series $\mathcal{H}^{(n+1)}_{k-\frac{1}{2}}$ (see (1.1) in §1 for the definition of $e^{(n)}_{k-\frac{1}{2},m}$). We remark $e^{(n)}_{k-\frac{1}{2},m} \in J^{(n)*}_{k-\frac{1}{2},m}$ for any integer n.

4.5 The map $\iota_{\mathcal{M}}$

We recall $\mathcal{M} = \begin{pmatrix} l & r/2 \\ r/2 & 1 \end{pmatrix} \in \mathrm{Sym}^+_2$. In this subsection we shall introduce a map

$$\iota_{\mathcal{M}} : H^{(n)}_{\mathcal{M}} \to \mathrm{Hol}(\mathfrak{H}_n \times \mathbb{C}^{(n,1)} \to \mathbb{C}),$$

where $H^{(n)}_{\mathcal{M}}$ is a subspace of holomorphic functions on $\mathfrak{H}_n \times \mathbb{C}^{(n,2)}$, which will be defined below, and where $\mathrm{Hol}(\mathfrak{H}_n \times \mathbb{C}^{(n,1)} \to \mathbb{C})$ denotes the space of all holomorphic functions on $\mathfrak{H}_n \times \mathbb{C}^{(n,1)}$. We will show that the restriction of $\iota_{\mathcal{M}}$ gives a linear isomorphism between $J_{k,\mathcal{M}}^{(n)*}$ and $J_{k-\frac{1}{2},m}^{(n)*}$ (cf. Lemma 4.2).

Let ψ be a holomorphic function on $\mathfrak{H}_n \times \mathbb{C}^{(n,2)}$. We assume that ψ has a Fourier expansion

$$\psi(\tau, z) \quad = \sum_{\substack{N \in Sym_n^*, R \in \mathbb{Z}^{(n,1)} \\ 4N - R\mathcal{M}^{-1\,t}R \geq 0}} A(N,R)\, e(N\tau +{}^t Rz)$$

for $(\tau, z) \in \mathfrak{H}_n \times \mathbb{C}^{(n,2)}$, and assume that ψ satisfies the following condition on the Fourier coefficients: if

$$\begin{pmatrix} N & \frac{1}{2}R \\ \frac{1}{2}{}^tR & \mathcal{M} \end{pmatrix} \quad = \quad \begin{pmatrix} N' & \frac{1}{2}R' \\ \frac{1}{2}{}^tR' & \mathcal{M} \end{pmatrix} \left[\begin{pmatrix} 1_n & \\ {}^tT & 1_2 \end{pmatrix} \right]$$

with some $T = \begin{pmatrix} 0, \lambda \end{pmatrix} \in \mathbb{Z}^{(n,2)}$ and some $\lambda \in \mathbb{Z}^{(n,1)}$, then $A(N,R) = A(N',R')$.

The symbol $H_{\mathcal{M}}^{(n)}$ denotes the $\mathbb{C}$-vector space consists of all holomorphic functions which satisfy the above condition.

We remark $J_{k,\mathcal{M}}^{(n)*} \subset J_{k,\mathcal{M}}^{(n)} \subset H_{\mathcal{M}}^{(n)}$ for any even integer k.

Now we shall define a map $\iota_{\mathcal{M}}$. For $\psi(\tau', z') = \sum A(N,R) e(N\tau' + R^t z') \in H_{\mathcal{M}}^{(n)}$ we define a holomorphic function $\iota_{\mathcal{M}}(\psi)$ on $\mathfrak{H}_n \times \mathbb{C}^{(n,1)}$ by

$$\iota_{\mathcal{M}}(\psi)(\tau, z) \quad := \sum_{\substack{M \in Sym_n^*,\ S \in \mathbb{Z}^{(n,1)} \\ 4Mm - S^tS \geq 0}} C(M,S) e(M\tau + S^t z)$$

for $(\tau, z) \in \mathfrak{H}_n \times \mathbb{C}^{(n,1)}$, where we define $C(M,S) := A(N,R)$ if there exist matrices $N \in \mathrm{Sym}_2^*$ and $R = (R_1, R_2) \in \mathbb{Z}^{(n,2)}$ $(R_1, R_2 \in \mathbb{Z}^{(n,1)})$ which satisfy

$$\begin{pmatrix} M & \frac{1}{2}S \\ \frac{1}{2}{}^tS & \det(2\mathcal{M}) \end{pmatrix} \quad = \quad 4 \begin{pmatrix} N & \frac{1}{2}R_1 \\ \frac{1}{2}{}^tR_1 & l \end{pmatrix} - \begin{pmatrix} R_2 \\ r \end{pmatrix} \begin{pmatrix} {}^tR_2, r \end{pmatrix},$$

$C(M,S) := 0$ otherwise. We remark that the identity

$$4 \begin{pmatrix} N & \frac{1}{2}R_1 \\ \frac{1}{2}{}^tR_1 & l \end{pmatrix} - \begin{pmatrix} R_2 \\ r \end{pmatrix} \begin{pmatrix} {}^tR_2, r \end{pmatrix} \quad = \quad 4 \begin{pmatrix} N & \frac{1}{2}R \\ \frac{1}{2}{}^tR & \mathcal{M} \end{pmatrix} \left[\begin{pmatrix} & 0 \\ 1_n & \vdots \\ & 0 \\ 0 \cdots 0 & 1 \\ -\frac{1}{2}{}^tR_2 & -\frac{1}{2}r \end{pmatrix} \right]$$

holds and remark that the coefficient $C(M,S)$ does not depend on the choice of the matrices N and R. The proof of these facts are as follows. The first fact of the identity follows from a straightforward calculation. As for the second fact, if

$$4 \begin{pmatrix} N & \frac{1}{2}R_1 \\ \frac{1}{2}{}^tR_1 & l \end{pmatrix} - \begin{pmatrix} R_2 \\ r \end{pmatrix} \begin{pmatrix} {}^tR_2, r \end{pmatrix} \quad = \quad 4 \begin{pmatrix} N' & \frac{1}{2}R_1' \\ \frac{1}{2}{}^tR_1' & l \end{pmatrix} - \begin{pmatrix} R_2' \\ r \end{pmatrix} \begin{pmatrix} {}^tR_2', r \end{pmatrix},$$

then $4N - R_2{}^tR_2 = 4N' - R_2'{}^tR_2'$. Hence $R_2{}^tR_2 \equiv R_2'{}^tR_2' \bmod 4$. Thus there exists a matrix $\lambda \in \mathbb{Z}^{(n,1)}$ such that $R_2' = R_2 + 2\lambda$. Therefore, by straightforward calculation we have

$$\begin{pmatrix} N & \frac{1}{2}R \\ \frac{1}{2}{}^tR & \mathcal{M} \end{pmatrix} = \begin{pmatrix} N' & \frac{1}{2}R' \\ \frac{1}{2}{}^tR' & \mathcal{M} \end{pmatrix} \left[\begin{pmatrix} 1_n & 0 \\ {}^tT & 1_2 \end{pmatrix} \right]$$

with $T = (0, \lambda)$, $R = (R_1, R_2)$ and $R' = (R'_1, R'_2)$. Because ψ belongs to $H_{\mathcal{M}}^{(n)}$, we have $A(N, R) = A(N', R')$. Hence the above definition of $C(M, S)$ is well-defined.

Lemma 4.2. *Let k be an even integer. We put $m = \det(2\mathcal{M})$. Then we have the commutative diagram:*

$$\begin{array}{ccc} J_{k,1}^{(n+1)} & \xrightarrow{\sigma} & M^+_{k-\frac{1}{2}}(\Gamma_0^{(n+1)}(4)) \\ {\scriptstyle FJ_{1,\mathcal{M}}}\downarrow & & \downarrow{\scriptstyle \widetilde{FJ}_m} \\ J_{k,\mathcal{M}}^{(n)*} & \xrightarrow{\iota_{\mathcal{M}}} & J_{k-\frac{1}{2},m}^{(n)*}, \end{array}$$

where two maps $FJ_{1,\mathcal{M}}$ and $\widetilde{FJ}_m$ have been introduced in §4.1 and §4.4. Moreover, the restriction of the linear map $\iota_{\mathcal{M}}$ on $J_{k,\mathcal{M}}^{(n)*}$ gives the bijection between $J_{k,\mathcal{M}}^{(n)*}$ and $J_{k-\frac{1}{2},m}^{(n)*}$.

Proof. Let $\psi \in J_{k,1}^{(n+1)}$ be a Jacobi form. Due to the definition of σ (cf. §4.3) and $\iota_{\mathcal{M}}$, it is not difficult to check the identity $\iota_M(FJ_{1,\mathcal{M}}(\psi)) = \widetilde{\mathrm{FJ}}_m(\sigma(\psi))$. Namely, we have the above commutative diagram.
Since the restriction of the map $\widetilde{\mathrm{FJ}}_m$ on $M^+_{k-\frac{1}{2}}(\Gamma_0^{(n+1)}(4))$ is surjective, and since σ is an isomorphism and since $\iota_M(FJ_{1,\mathcal{M}}(\psi)) = \widetilde{\mathrm{FJ}}_m(\sigma(\psi))$, the restricted map $\iota_{\mathcal{M}}|_{J_{k,\mathcal{M}}^{(n)*}} : J_{k,\mathcal{M}}^{(n)*} \to J_{k-\frac{1}{2},m}^{(n)*}$ is surjective. The injectivity of the restricted map $\iota_{\mathcal{M}}|_{J_{k,\mathcal{M}}^{(n)*}}$ follows directly from the definition of the map $\iota_{\mathcal{M}}$. $\square$

4.6 Compatibility between index-shift maps and $\iota_{\mathcal{M}}$

In this subsection we shall show a compatibility between the map $\iota_{\mathcal{M}}$ and some index-shift maps.
For function ψ on $\mathfrak{H}_n \times \mathbb{C}^{(n,2)}$ and for $L \in \mathbb{Z}^{(2,2)}$ we define the function $\psi|U_L$ on $\mathfrak{H}_n \times \mathbb{C}^{(n,2)}$ by

$$(\psi|U_L)(\tau, z) := \psi(\tau, z\,{}^tL).$$

It is not difficult to check that if ψ belongs to $J_{k,\mathcal{M}}^{(n)}$, then $\psi|U_L$ belongs to $J_{k,\mathcal{M}[L]}^{(n)}$.

For function ϕ on $\mathfrak{H}_n \times \mathbb{C}^{(n,1)}$ and for integer a we define the function $\phi|U_a$ on $\mathfrak{H}_n \times \mathbb{C}^{(n,1)}$ by

$$(\phi|U_a)(\tau, z) \quad := \quad \phi(\tau, az).$$

We have $\phi|U_a \in J^{(n)}_{k-\frac{1}{2}, ma^2}$ if $\phi \in J^{(n)}_{k-\frac{1}{2}, m}$.

Proposition 4.3. *For any $\psi \in J^{(n)*}_{k,\mathcal{M}}$ and for any $L = \left(\begin{smallmatrix} a & \\ b & 1 \end{smallmatrix}\right) \in \mathbb{Z}^{(2,2)}$ we obtain*

$$\iota_{\mathcal{M}[L]}(\psi|U_L) \quad = \quad \iota_{\mathcal{M}}(\psi)|U_a.$$

In particular, for any prime p we have $\iota_{\mathcal{M}[\left(\begin{smallmatrix} p & \\ & 1 \end{smallmatrix}\right)]}\left(\psi \,\middle|\, U_{\left(\begin{smallmatrix} p & \\ & 1 \end{smallmatrix}\right)}\right) = \iota_{\mathcal{M}}(\psi)|U_p$.

Proof. We put $m = \det(2\mathcal{M})$. Let $\psi(\tau, z') = \sum_{\substack{N \in Sym_n^*,\ R \in \mathbb{Z}^{(n,2)} \\ 4N - R\mathcal{M}^{-1}\,{}^tR \geq 0}} A(N,R) e(N\tau + R\,{}^tz')$ be a Fourier expansion of ψ. Let

$$\begin{aligned}
\iota_{\mathcal{M}}(\psi)(\tau, z) &= \sum_{\substack{M \in Sym_n^*,\ S \in \mathbb{Z}^{(n,1)} \\ 4Mm - S\,{}^tS \geq 0}} C(M,S)\, e(M\tau + S\,{}^tz), \\
\iota_{\mathcal{M}[L]}(\psi|U_L)(\tau, z) &= \sum_{\substack{M \in Sym_n^*,\ S \in \mathbb{Z}^{(n,1)} \\ 4Mma^2 - S\,{}^tS \geq 0}} C_1(M,S)\, e(M\tau + S\,{}^tz)
\end{aligned}$$

and

$$(\iota_{\mathcal{M}}(\psi)|U_a)(\tau, z) = \sum_{\substack{M \in Sym_n^*,\ S \in \mathbb{Z}^{(n,1)} \\ 4Mma^2 - S\,{}^tS \geq 0}} C_2(M,S)\, e(M\tau + S\,{}^tz)$$

be Fourier expansions. It is enough to show $C_1(M,S) = C_2(M,S)$.
We have $C_2(M,S) = C(M, a^{-1}S)$. Moreover, we obtain $C_1(M,S) = A(N, RL^{-1})$ with $N \in \mathrm{Sym}_n^*$ and $R \in \mathbb{Z}^{(n,2)}$ which satisfy

$$\begin{pmatrix} M & \frac{1}{2}S \\ \frac{1}{2}{}^tS & ma^2 \end{pmatrix} \quad = \quad 4 \begin{pmatrix} N & \frac{1}{2}R \\ \frac{1}{2}{}^tR & \mathcal{M}[L] \end{pmatrix} \left[\begin{pmatrix} & 0 \\ 1_n & \vdots \\ & 0 \\ 0 \cdots 0 & 1 \\ -\frac{1}{2}{}^t(R\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)) & -\frac{1}{2}ra - b \end{pmatrix} \right].$$

For the above matrices N, R, M and S we have

$$
\begin{aligned}
&\begin{pmatrix} M & \frac{1}{2}a^{-1}S \\ \frac{1}{2}a^{-1\,t}S & m \end{pmatrix} \\
&= 4\begin{pmatrix} N & \frac{1}{2}R \\ \frac{1}{2}{}^tR & \mathcal{M}[L] \end{pmatrix}\left[\begin{pmatrix} & & 0 \\ & 1_n & \vdots \\ & & 0 \\ & 0\cdots 0 & 1 \\ -\frac{1}{2}{}^t(R\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)) & & -\frac{1}{2}ra-b \end{pmatrix}\begin{pmatrix} & & 0 \\ & 1_n & \vdots \\ & & 0 \\ 0\cdots 0 & & a^{-1} \end{pmatrix}\right] \\
&= 4\begin{pmatrix} N & \frac{1}{2}RL^{-1} \\ \frac{1}{2}{}^t(RL^{-1}) & \mathcal{M} \end{pmatrix}\left[\begin{pmatrix} & & 0 \\ & 1_n & \vdots \\ & & 0 \\ & 0\cdots 0 & 1 \\ -\frac{1}{2}{}^t(R\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)) & & -\frac{1}{2}r \end{pmatrix}\right] \\
&= 4\begin{pmatrix} N & \frac{1}{2}RL^{-1} \\ \frac{1}{2}{}^t(RL^{-1}) & \mathcal{M} \end{pmatrix}\left[\begin{pmatrix} & & 0 \\ & 1_n & \vdots \\ & & 0 \\ & 0\cdots 0 & 1 \\ -\frac{1}{2}{}^t(RL^{-1}\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)) & & -\frac{1}{2}r \end{pmatrix}\right].
\end{aligned}
$$

Thus $C_2(M,S) = C(M,a^{-1}S) = A(N,RL^{-1}) = C_1(M,S)$. □

PROPOSITION 4.4. *For odd prime p and for $0 \le \alpha \le n$, let $\tilde{V}_{\alpha,n-\alpha}(p^2)$ and $V_{\alpha,n-\alpha}(p^2)$ be index-shift maps defined in §2.7. Then, for any $\psi \in J_{k,\mathcal{M}}^{(n)*}$ we have*

$$\iota_{\mathcal{M}}(\psi)|\tilde{V}_{\alpha,n-\alpha}(p^2) \quad = \quad p^{k(2n+1)-n(n+\frac{7}{2})+\frac{1}{2}\alpha}\ \iota_{\mathcal{M}[\left(\begin{smallmatrix}p & \\ & 1\end{smallmatrix}\right)]}(\psi|V_{\alpha,n-\alpha}(p^2)). \tag{4.3}$$

Proof. The proof is similar to the case of Jacobi forms of index 1 (cf. [Ib 92, Theorem 2]). However, we remark that the maps $\tilde{V}_{\alpha,n-\alpha}(p^2)$ and $V_{\alpha,n-\alpha}(p^2)$ in the present article change the indices of Jacobi forms.

To prove this proposition, we compare the Fourier coefficients of the both sides of (4.3). Let

$$
\begin{aligned}
\psi(\tau,z') &= \sum_{N,R} A_1(N,R)e(N\tau + R\,{}^tz'), \\
(\psi|V_{\alpha,n-\alpha}(p^2))(\tau,z') &= \sum_{N,R} A_2(N,R)e(N\tau + R\,{}^tz'), \\
(\iota_{\mathcal{M}}(\psi))(\tau,z) &= \sum_{M,S} C_1(M,S)e(M\tau + S\,{}^tz)
\end{aligned}
$$

and

$$(\iota_{\mathcal{M}}(\psi)|\tilde{V}_{\alpha,n-\alpha}(p^2))(\tau,z) \quad = \quad \sum_{M,S} C_2(M,S)e(M\tau + S\,{}^tz)$$

be Fourier expansions, where $\tau \in \mathfrak{H}_n$, $z' \in \mathbb{C}^{(n,2)}$ and $z \in \mathbb{C}^{(n,1)}$. For the sake of simplicity we put $U = \begin{pmatrix} p^2 & \\ & p \end{pmatrix}$. Then

$$\begin{aligned}
&\psi|V_{\alpha,n-\alpha}(p^2) \\
&= \sum_{\begin{pmatrix} p^2\,{}^tD^{-1} & B \\ 0_n & D \end{pmatrix}} \sum_{\lambda_2,\mu_2 \in (\mathbb{Z}/p\mathbb{Z})^{(n,1)}} \\
&\qquad \times \psi|_{k,\mathcal{M}} \left(\begin{pmatrix} p^2\,{}^tD^{-1} & B \\ 0_n & D \end{pmatrix} \times \begin{pmatrix} U & \\ & p^2U^{-1} \end{pmatrix}, [((0,\lambda_2),(0,\mu_2)),0_2] \right) \\
&= \sum_{\begin{pmatrix} p^2\,{}^tD^{-1} & B \\ 0_n & D \end{pmatrix}} \sum_{\lambda_2,\mu_2 \in (\mathbb{Z}/p\mathbb{Z})^{(n,1)}} \sum_{N,R} A(N,R) \\
&\qquad \times e(N\tau + R\,{}^tz)|_{k,\mathcal{M}} \left(\begin{pmatrix} p^2\,{}^tD^{-1} & B \\ 0_n & D \end{pmatrix} \times \begin{pmatrix} U & \\ & p^2U^{-1} \end{pmatrix}, [((0,\lambda_2),(0,\mu_2)),0_2] \right),
\end{aligned}$$

where $\begin{pmatrix} p^2\,{}^tD^{-1} & B \\ 0_n & D \end{pmatrix}$ runs over a set of all representatives of

$$\Gamma_n \backslash \Gamma_n \mathrm{diag}(1_\alpha, p1_{n-\alpha}, p^2 1_\alpha, p1_{n-\alpha})\Gamma_n,$$

and where the slash operator $|_{k,\mathcal{M}}$ is defined in §2.4.
We put $\lambda = (0,\lambda_2)$, $\mu = (0,\mu_2) \in \mathbb{Z}^{(n,2)}$, then we obtain

$$\begin{aligned}
&e(N\tau + R\,{}^tz)|_{k,\mathcal{M}} \left(\begin{pmatrix} p^2\,{}^tD^{-1} & B \\ 0_n & D \end{pmatrix} \times \begin{pmatrix} U & \\ & p^2U^{-1} \end{pmatrix}, [(\lambda,\mu),0_2] \right) \\
&= p^{-k}\det(D)^{-k} e(\hat{N}\tau + \hat{R}\,{}^tz + NBD^{-1} + RU\,{}^t\mu D^{-1}),
\end{aligned}$$

where

$$\hat{N} \;=\; p^2 D^{-1} N\,{}^tD^{-1} + D^{-1}RU\,{}^t\lambda + \frac{1}{p^2}\lambda U\mathcal{M}U\,{}^t\lambda$$

and

$$\hat{R} \;=\; D^{-1}RU + \frac{2}{p^2}\lambda U\mathcal{M}U.$$

Thus

$$N \;=\; \frac{1}{p^2} D \left(\left(\hat{N} - \frac{1}{4}\hat{R}_2\,{}^t\hat{R}_2 \right) + \frac{1}{4}(\hat{R}_2 - 2\lambda_2)^t(\hat{R}_2 - 2\lambda_2) \right) {}^tD$$

and

$$R \;=\; D\left(\hat{R} - \frac{2}{p^2}\lambda U\mathcal{M}U \right) U^{-1},$$

where $\hat{R}_2 = \hat{R}\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$. Hence, for any $\hat{N} \in \mathrm{Sym}_n^*$ and for any $\hat{R} \in \mathbb{Z}^{(n,2)}$, we have

$$\begin{aligned}
&A_2(\hat{N},\hat{R})\\
&= p^{-k} \sum_{\left(\begin{smallmatrix} p^{2\,t}D^{-1} & B \\ 0_n & D\end{smallmatrix}\right)} \det(D)^{-k} \sum_{\lambda_2 \in (\mathbb{Z}/p\mathbb{Z})^{(n,1)}} \sum_{\mu_2 \in (\mathbb{Z}/p\mathbb{Z})^{(n,1)}} A_1(N,R)\\
&\qquad \times e(NBD^{-1} + RU^t(0,\mu_2)D^{-1})\\
&= p^{-k+n} \sum_{\left(\begin{smallmatrix} p^{2\,t}D^{-1} & B \\ 0_n & D\end{smallmatrix}\right)} \det(D)^{-k} \sum_{\lambda_2 \in (\mathbb{Z}/p\mathbb{Z})^{(n,1)}} A_1(N,R)\, e(NBD^{-1}),
\end{aligned}$$

where N and R are the same symbols as above, which are determined by $\hat{N}$, $\hat{R}$ and λ_2, and where $\left(\begin{smallmatrix} p^{2\,t}D^{-1} & B \\ 0_n & D\end{smallmatrix}\right)$ runs over a complete set of representatives of

$$\Gamma_n \backslash \Gamma_n \mathrm{diag}(1_\alpha, p1_{n-\alpha}, p^2 1_\alpha, p1_{n-\alpha})\Gamma_n.$$

On the RHS of the above first identity the matrix $D^{-1}RU$ belongs to $\mathbb{Z}^{(n,2)}$, since $\hat{R} \in \mathbb{Z}^{(n,2)}$. We remark that $A_1(N,R) = 0$ unless $N \in \mathrm{Sym}_n^*$ and $R \in \mathbb{Z}^{(n,2)}$.
Due to the definition of $\iota_{\mathcal{M}}$, for $N \in \mathrm{Sym}_n^*$ and $R \in \mathbb{Z}^{(n,2)}$ we have the identity

$$A_1(N,R) \quad = \quad C_1(4N - R\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)^t (R\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)), 4R\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right) - 2rR\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)).$$

Here

$$4N - R\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)^t(R\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)) \quad = \quad \frac{1}{p^2} D\left(4\hat{N} - \hat{R}_2{}^t\hat{R}_2\right){}^tD$$

and

$$4R\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right) - 2rR\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right) \quad = \quad \frac{1}{p^2}D(4\hat{R}\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right) - 2rp\hat{R}_2).$$

Hence we have

$$\begin{aligned}
&A_2(\hat{N},\hat{R}) \qquad\qquad (4.4)\\
= \;& p^{-k+n} \sum_{\left(\begin{smallmatrix} p^{2\,t}D^{-1} & B \\ 0_n & D\end{smallmatrix}\right)} \det(D)^{-k}\\
&\times C_1\left(\frac{1}{p^2}D\left(4\hat{N} - \hat{R}_2{}^t\hat{R}_2\right){}^tD, \frac{1}{p^2}D(4\hat{R}\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right) - 2rp\hat{R}_2)\right)\\
&\times e\left(\frac{1}{p^2}\left(\hat{N} - \frac{1}{4}\hat{R}_2{}^t\hat{R}_2\right){}^tDB\right) \sum_{\lambda_2} e\left(\frac{1}{4p^2}(\hat{R}_2 - 2\lambda_2)^t(\hat{R}_2 - 2\lambda_2)^tDB\right),
\end{aligned}$$

where λ_2 runs over a complete set of representatives of $(\mathbb{Z}/p\mathbb{Z})^{(n,1)}$ such that

$$D\left(\hat{R} - \frac{2}{p^2}(0,\lambda_2)U\mathcal{M}U\right)U^{-1} \in \mathbb{Z}^{(n,2)}.$$

Let $\mathfrak{S}_\alpha$ be a complete set of representative of $\Gamma_n\backslash\Gamma_n \begin{pmatrix} 1_\alpha & & & \\ & p1_{n-\alpha} & & \\ & & p^2 1_\alpha & \\ & & & p1_{n-\alpha}\end{pmatrix}\Gamma_n$. Now we quote a complete set of representatives $\mathfrak{S}_\alpha$ from [Zh 84]. We put

$$\delta_{i,j} := \mathrm{diag}(1_i, p1_{j-i}, p^2 1_{n-j})$$

for $0 \le i \le j \le n$. We set

$$\mathfrak{S}_\alpha \quad := \quad \left\{ \begin{pmatrix} p^2{\delta_{i,j}}^{-1} & b_0 \\ 0_n & \delta_{i,j}\end{pmatrix}\begin{pmatrix} {}^t u^{-1} & 0_n \\ 0_n & u\end{pmatrix} \middle|\, i,j,b_0,u \right\},$$

where i and j run over all non-negative integers such that $j-i-n+\alpha \ge 0$, and where u runs over a complete set of representatives of $(\delta_{i,j}^{-1}\mathrm{GL}_n(\mathbb{Z})\delta_{i,j} \cap \mathrm{GL}_n(\mathbb{Z}))\backslash \mathrm{GL}_n(\mathbb{Z})$, and b_0 runs over all matrices in the set

$$\mathfrak{T} := \left\{ \begin{pmatrix} 0_i & 0 & 0 \\ 0 & a_1 & pb_1 \\ 0 & {}^t b_1 & b_2 \end{pmatrix} \middle|\, \begin{matrix} b_1 \in (\mathbb{Z}/p\mathbb{Z})^{(j-i,n-j)}, b_2 = {}^t b_2 \in (\mathbb{Z}/p^2\mathbb{Z})^{(n-j,n-j)}, \\ a_1 = {}^t a_1 \in (\mathbb{Z}/p\mathbb{Z})^{(j-i,j-i)}, \mathrm{rank}_p(a_1) = j-i-n+\alpha \end{matrix} \right\}.$$

For a matrix $g = \begin{pmatrix} p^{2t}D^{-1} & B \\ 0_n & D\end{pmatrix} = \begin{pmatrix} p^2{\delta_{i,j}}^{-1} & b_0 \\ 0_n & \delta_{i,j}\end{pmatrix}\begin{pmatrix} {}^t u^{-1} & 0_n \\ 0_n & u\end{pmatrix} \in \mathfrak{S}_\alpha$ with a matrix $b_0 = \begin{pmatrix} 0_i & 0 & 0 \\ 0 & a_1 & pb_1 \\ 0 & {}^t b_1 & b_2 \end{pmatrix} \in \mathfrak{T}$, we define $\varepsilon(g) := \left(\frac{-4}{p}\right)^{rank_p(a_1)/2}\left(\frac{\det a_1'}{p}\right)$, where $a_1' \in \mathrm{GL}_{j-i-n+\alpha}(\mathbb{Z}/p\mathbb{Z})$ is a matrix such that $a_1 \equiv \begin{pmatrix} a_1' & 0 \\ 0 & 0_{n-\alpha}\end{pmatrix}[v] \mod p$ with some $v \in \mathrm{GL}_{j-i}(\mathbb{Z})$. Under the assumption

$$\frac{1}{p^2}D(4\hat{R}\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right) - 2rp\hat{R}_2) \in \mathbb{Z}^{(n,1)},$$

the condition $D(\hat{R} - 2p^{-2}(0,\lambda_2)U\mathcal{M}U)U^{-1} \in \mathbb{Z}^{(n,2)}$ is equivalent to the condition

$$u(\hat{R}_2 - 2\lambda_2) \in \begin{pmatrix} p1_i & 0 \\ 0 & 1_{n-i}\end{pmatrix}\mathbb{Z}^{(n,1)}.$$

Hence the last summation in (4.4) is

$$\begin{aligned} & \sum_{\lambda_2} e\left(\frac{1}{4p^2}(\hat{R}_2 - 2\lambda_2)^t(\hat{R}_2 - 2\lambda_2)^t DB\right) \\ = \quad & p^{n-j} \sum_{\lambda' \in (\mathbb{Z}/p\mathbb{Z})^{(j-i,1)}} e\left(\frac{1}{p}{}^t\lambda' a_1 \lambda'\right) \\ = \quad & p^{n-i-rank_p(a_1)}\left(\left(\frac{-4}{p}\right)p\right)^{rank_p(a_1)/2}\left(\frac{\det a_1'}{p}\right) \\ = \quad & p^{n-i-\frac{rank_p(a_1)}{2}}\varepsilon(g) \\ = \quad & p^{n+(n-i-j-\alpha)/2}\varepsilon(g). \end{aligned}$$

Thus (4.4) is

$$\begin{aligned}
&A_2(\hat{N},\hat{R})\\
&= p^{-k+2n}\sum_{g} p^{-k(2n-i-j)+(n-i-j-\alpha)/2}\varepsilon(g)\, e\Big(p^{-2}\left(4\hat{N}-\hat{R}_2{}^t\hat{R}_2\right){}^tDB\Big)\\
&\quad\times C_1\Big(p^{-2}D(4\hat{N}-\hat{R}_2{}^t\hat{R}_2)^tD,\ p^{-2}D(4\hat{R}\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)-2rp\hat{R}_2)\Big),
\end{aligned}$$

where $g = \begin{pmatrix} p^{2t}D^{-1} & B \\ 0_n & D\end{pmatrix} = \begin{pmatrix} p^2\delta_{i,j}{}^{-1} & b_0 \\ 0_n & \delta_{i,j}\end{pmatrix}\begin{pmatrix} {}^tu^{-1} & 0_n \\ 0_n & u\end{pmatrix}$ runs over all elements in the set $\mathfrak{S}_\alpha$.

Now we shall express $C_2(M,S)$ as a linear combination of Fourier coefficients $C_1(M,S)$ of $\iota_M(\psi)$. For $Y = (\mathrm{diag}(1_\alpha, p1_{n-\alpha}, p^2 1_\alpha, p1_{n-\alpha}), p^{\alpha/2}) \in \widetilde{\mathrm{GSp}_n^+}(\mathbb{Z})$ a complete set of representatives of $\Gamma_0^{(n)}(4)^*\backslash\Gamma_0^{(n)}(4)^*Y\Gamma_0^{(n)}(4)^*$ is given by elements

$$\widetilde{g} \quad=\quad (g,\varepsilon(g)p^{(n-i-j)/2}) \in \widetilde{\mathrm{GSp}_n^+}(\mathbb{Z}),$$

where g runs over all elements in the set $\mathfrak{S}_\alpha$, and $\varepsilon(g)$ is defined as above (cf. [Zh 84, Lemma 3.2]). Hence

$$\begin{aligned}
&(\iota_{\mathcal{M}}(\psi)|\tilde{V}_{\alpha,n-\alpha}(p^2))(\tau,z)\\
&= p^{n(2k-1)/2-n(n+1)}\sum_{M,S}\sum_{\widetilde{g}} p^{(-k+1/2)(n-i-j)}\varepsilon(g)\, C_1(M,S)\\
&\qquad\qquad\qquad\times e(M(p^{2t}D^{-1}\tau+B)D^{-1}+p^2S^tzD^{-1})\\
&= p^{n(2k-1)/2-n(n+1)}\sum_{\hat{M},\hat{S}}\sum_{g\in\mathfrak{S}_\alpha} p^{(-k+1/2)(n-i-j)}\varepsilon(g)\, C_1(p^{-2}D\hat{M}^tD, p^{-2}D\hat{S})\\
&\qquad\qquad\qquad\times e(\hat{M}\tau+\hat{S}^tz+p^{-2}\hat{M}^tDB).
\end{aligned}$$

Thus

$$\begin{aligned}
&C_2(\hat{M},\hat{S})\\
&= \sum_{g} p^{-n(n+1)+(k-1/2)(i+j)}\varepsilon(g)\, C_1(p^{-2}D\hat{M}^tD, p^{-2}D\hat{S})\, e(p^{-2}\hat{M}^tDB).
\end{aligned}$$

Now we put $\hat{M} = 4\hat{N}-\hat{R}_2{}^t\hat{R}_2$ and $\hat{S} = 4\hat{R}\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)-2rp\hat{R}_2$, then

$$C_2(4\hat{N}-\hat{R}_2{}^t\hat{R}_2, 4\hat{R}\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)-2rp\hat{R}_2) \quad=\quad p^{2nk+k-n^2-\frac{7}{2}n+\frac{1}{2}\alpha}A_2(\hat{N},\hat{R}).$$

The proposition follows from this identity. □

4.7 Index-shift maps at $p=2$

For $p=2$ we define the map

$$\tilde{V}_{\alpha,n-\alpha}(4) \ :\ J_{k-\frac{1}{2},m}^{(n)*} \to \mathrm{Hol}(\mathfrak{H}_n\times\mathbb{C}^{(n,1)}\to\mathbb{C})$$

through an analogue of the identity (4.3), it means that we define

$$\phi|\tilde{V}_{\alpha,n-\alpha}(4) \quad := \quad 2^{k(2n+1)-n(n+\frac{7}{2})+\frac{1}{2}\alpha}\ \iota_{\mathcal{M}[\left(\begin{smallmatrix}2 & \\ & 1\end{smallmatrix}\right)]}(\psi|V_{\alpha,n-\alpha}(4))$$

for any $\phi \in J^{(n)*}_{k-\frac{1}{2},m}$, and where $\psi \in J^{(n)*}_{k,\mathcal{M}}$ is the Jacobi form which satisfies $\iota_{\mathcal{M}}(\psi) = \phi$. Here the map $V_{\alpha,n-\alpha}(4)$ is defined in §2.7 and the map $\iota_{\mathcal{M}}$ is defined in §4.5.

5 Action of index-shift maps on Jacobi-Eisenstein series

In this section we fix a positive definite half-integral symmetric matrix $\mathcal{M} \in \mathrm{Sym}_2^+$ and we assume that the right-lower part of $\mathcal{M}$ is 1, it means $\mathcal{M} = \begin{pmatrix} * & * \\ * & 1\end{pmatrix}$.

The purpose of this section is to show that the form $E^{(n)}_{k,\mathcal{M}}|V_{\alpha,n-\alpha}(p^2)$ is written as a linear combination of three forms $E^{(n)}_{k,\mathcal{M}\left[\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)\right]}$, $E^{(n)}_{k,\mathcal{M}}|U_{\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)}$ and $E^{(n)}_{k,\mathcal{M}\left[X^{-1}\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)^{-1}\right]}|U_{\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)X\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)}$, where $E^{(n)}_{k,\mathcal{M}}$ is the Jacobi-Eisenstein series of index $\mathcal{M}$ (cf. §3), and where $V_{\alpha,n-\alpha}(p^2)$ and $U_{\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)}$ are index-shift maps (cf. §2.7 and §4.6). Here $X = \left(\begin{smallmatrix}1&0\\x&1\end{smallmatrix}\right)$ is a matrix.

First we will calculate functions $K^{\beta}_{i,j}$ (cf. Lemma 5.2) which appear in an expression of $E^{(n)}_{k,\mathcal{M}}|V_{\alpha,n-\alpha}(p^2)$, and after that, we will express $E^{(n)}_{k,\mathcal{M}}|V_{\alpha,n-\alpha}(p^2)$ as a summation of functions $\tilde{K}^{\beta}_{i,j}$ (cf. Proposition 5.3).

The calculation in this section is an analogue to the one given in [Yk 89] for the case of index $\mathcal{M} = 1$. However, we need to modify his calculation for Jacobi-Eisenstein series $E^{(n)}_{k,1}$ of index 1 to our case for $E^{(n)}_{k,\mathcal{M}}$ with $\mathcal{M} = \left(\begin{smallmatrix}*&*\\*&1\end{smallmatrix}\right) \in \mathrm{Sym}_2^+$. This calculation is not obvious, since we need to calculate the action of the matrices of type $[((0,u_2),(0,v_2)),0_2]$.

5.1 The function $K^{\beta}_{i,j}$

The purpose of this subsection is to introduce a function $K^{\beta}_{i,j}$ and to express $E^{(n)}_{k,\mathcal{M}}|V_{\alpha,n-\alpha}(p^2)$ as a summation over $K^{\beta}_{i,j}$. Moreover, we shall calculate $K^{\beta}_{i,j}$ explicitly (cf. Lemma 5.2).

We put $\delta_{i,j} := \mathrm{diag}(1_i, p1_{j-i}, p^2 1_{n-j})$. For $x = \mathrm{diag}(0_i, x', 0_{n-i-j})$ with $x' = {}^t x' \in \mathbb{Z}^{(j-i,j-i)}$, we set $\delta_{i,j}(x) := \begin{pmatrix} p^2\delta_{i,j}^{-1} & x \\ 0 & \delta_{i,j}\end{pmatrix}$ and $\Gamma(\delta_{i,j}(x)) := \Gamma_n \cap \delta_{i,j}(x)^{-1}\Gamma^{(n)}_{\infty}\delta_{i,j}(x)$.

For $x = \mathrm{diag}(0_i, x', 0_{n-i-j})$ and for $y = \mathrm{diag}(0_i, y', 0_{n-i-j})$ with $x' = {}^t x', y' = {}^t y' \in \mathbb{Z}^{(j-i,j-i)}$, following [Yk 89] we say that x and y are *equivalent*, if there exists a matrix $u \in \mathrm{GL}_n(\mathbb{Z}) \cap \delta_{i,j}\mathrm{GL}_n(\mathbb{Z})\delta_{i,j}^{-1}$ which has a form $u = \left(\begin{smallmatrix} u_1 & * & * \\ * & u_2 & * \\ * & * & u_3\end{smallmatrix}\right)$

satisfying $x' \equiv u_2\, y'\, {}^t u_2 \mod p$, where $u_2 \in \mathbb{Z}^{(j-i,j-i)}$, $u_1 \in \mathbb{Z}^{(i,i)}$ and $u_3 \in \mathbb{Z}^{(n-j,n-j)}$.
We denote by $[x]$ the equivalence class of x. We quote the following lemma from [Yk 89].

LEMMA 5.1. *The double coset* $\Gamma_n diag(1_\alpha, p1_{n-\alpha}, p^2 1_\alpha, p1_{n-\alpha})\Gamma_n$ *is written as a disjoint union*

$$\Gamma_n \begin{pmatrix} 1_\alpha &&& \\ & p1_{n-\alpha} && \\ && p^2 1_\alpha & \\ &&& p1_{n-\alpha} \end{pmatrix} \Gamma_n = \bigcup_{\substack{i,j \\ 0\le i\le j\le n}} \bigcup_{[x]} \Gamma_\infty^{(n)} \delta_{i,j}(x)\Gamma_n,$$

where $[x]$ *runs over all equivalence classes which satisfy* $rank_p(x) = j-i-n+\alpha \geq 0$.

Proof. The reader is referred to [Yk 89, Corollary 2.2]. $\square$

We put $U := \begin{pmatrix} p^2 & 0 \\ 0 & p \end{pmatrix}$. By the definition of index-shift map $V_{\alpha,n-\alpha}(p^2)$ and of the Jacobi-Eisenstein series $E_{k,\mathcal{M}}^{(n)}$, we have

$$\begin{aligned}
&E_{k,\mathcal{M}}^{(n)}|V_{\alpha,n-\alpha}(p^2) \\
&= \sum_{u,v\in\mathbb{Z}^{(n,1)}} \sum_{M'\in\Gamma_n\backslash\Gamma_n diag(1_\alpha,p1_{n-\alpha},p^2 1_\alpha,p1_{n-\alpha})\Gamma_n} \sum_{M\in\Gamma_\infty^{(n)}\backslash\Gamma_n} \sum_{\lambda\in\mathbb{Z}^{(n,2)}} \\
&\qquad \times 1|_{k,\mathcal{M}}([(\lambda,0),0_2], MM' \times \left(\begin{smallmatrix} U & 0 \\ 0 & p^2 U^{-1}\end{smallmatrix}\right))|_{k,\mathcal{M}[\left(\begin{smallmatrix} p & 0 \\ 0 & 1\end{smallmatrix}\right)]}[((0,u),(0,v)),0_2] \\
&= \sum_{u,v\in\mathbb{Z}^{(n,1)}} \sum_{M\in\Gamma_\infty^{(n)}\backslash\Gamma_n diag(1_\alpha,p1_{n-\alpha},p^2 1_\alpha,p1_{n-\alpha})\Gamma_n} \sum_{\lambda\in\mathbb{Z}^{(n,2)}} \\
&\qquad \times 1|_{k,\mathcal{M}}([(\lambda,0),0_2], M \times \left(\begin{smallmatrix} U & 0 \\ 0 & p^2 U^{-1}\end{smallmatrix}\right))|_{k,\mathcal{M}[\left(\begin{smallmatrix} p & 0 \\ 0 & 1\end{smallmatrix}\right)]}[((0,u),(0,v)),0_2].
\end{aligned}$$

Hence, due to Lemma 5.1, we have

$$\begin{aligned}
&E_{k,\mathcal{M}}^{(n)}|V_{\alpha,n-\alpha}(p^2) \\
&= \sum_{u,v\in\mathbb{Z}^{(n,1)}} \sum_{\substack{i,j \\ 0\le i\le j\le n}} \sum_{\substack{[x] \\ rank_p(x)=j-i-n+\alpha}} \sum_{M\in\Gamma_\infty^{(n)}\backslash\delta_{i,j}(x)\Gamma_n} \sum_{\lambda\in\mathbb{Z}^{(n,2)}} \\
&\qquad \times 1|_{k,\mathcal{M}}([(\lambda,0),0_2], M \times \left(\begin{smallmatrix} U & 0 \\ 0 & p^2 U^{-1}\end{smallmatrix}\right))|_{k,\mathcal{M}[\left(\begin{smallmatrix} p & 0 \\ 0 & 1\end{smallmatrix}\right)]}[((0,u),(0,v)),0_2] \\
&= \sum_{u,v\in\mathbb{Z}^{(n,1)}} \sum_{\substack{i,j \\ 0\le i\le j\le n}} \sum_{\substack{[x] \\ rank_p(x)=j-i-n+\alpha}} \sum_{M\in\Gamma(\delta_{i,j}(x))\backslash\Gamma_n} \sum_{\lambda\in\mathbb{Z}^{(n,2)}} \\
&\quad \times 1|_{k,\mathcal{M}}([(\lambda,0),0_2], \delta_{i,j}(x)M \times \left(\begin{smallmatrix} U & 0 \\ 0 & p^2 U^{-1}\end{smallmatrix}\right))|_{k,\mathcal{M}[\left(\begin{smallmatrix} p & 0 \\ 0 & 1\end{smallmatrix}\right)]}[((0,u),(0,v)),0_2].
\end{aligned}$$

For $\beta \le j-i$ we define a function

$$\begin{aligned}
&K_{i,j}^{\beta}(\tau,z)\\
&:= K_{i,j,\mathcal{M},p}^{\beta}(\tau,z)\\
&= \sum_{\substack{[x]\\ rank_p(x)=\beta}} \sum_{M\in\Gamma(\delta_{i,j}(x))\backslash\Gamma_n}\\
&\quad\times \sum_{\lambda\in\mathbb{Z}^{(n,2)}} \left\{1|_{k,\mathcal{M}}([(\lambda,0),0_2],\delta_{i,j}(x)M\times\left(\begin{smallmatrix}U & 0\\ 0 & p^2\,{}^tU^{-1}\end{smallmatrix}\right))\right\}(\tau,z).
\end{aligned}$$

Then we obtain

$$E_{k,\mathcal{M}}^{(n)}|V_{\alpha,n-\alpha}(p^2) \;=\; \sum_{\substack{i,j\\ 0\le i\le j\le n}} \sum_{u,v\in\mathbb{Z}^{(n,1)}} K_{i,j}^{\alpha-i-n+j}|_{k,\mathcal{M}[\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)]}[((0,u),(0,v)),0_2].$$

We define

$$L_{i,j} := L_{i,j,\mathcal{M},p} = \left\{ \begin{pmatrix}\lambda_1\\ \lambda_2\\ \lambda_3\end{pmatrix} \;\middle|\; \begin{array}{c}\lambda_1\in(p\mathbb{Z})^{(i,2)},\ \lambda_2\in\mathbb{Z}^{(j-i,2)},\ \lambda_3\in(p^{-1}\mathbb{Z})^{(n-j,2)}\\ 2\lambda_2\mathcal{M}^t\lambda_3\in\mathbb{Z}^{(j-i,n-j)},\ \lambda_3\mathcal{M}^t\lambda_3\in\mathbb{Z}^{(n-j,n-j)}\end{array}\right\}.$$

Moreover, we define a subgroup $\Gamma(\delta_{i,j})$ of $\Gamma_\infty^{(n)}$ by

$$\Gamma(\delta_{i,j}) \;:=\; \left\{\begin{pmatrix}A & B\\ 0_n & {}^tA^{-1}\end{pmatrix}\in\Gamma_\infty^{(n)} \;\middle|\; A\in\delta_{i,j}\mathrm{GL}_n(\mathbb{Z})\delta_{i,j}^{-1}\right\}.$$

Lemma 5.2. *Let $K_{i,j}^{\beta}$ be as above. We obtain*

$$\begin{aligned}
&K_{i,j}^{\beta}(\tau,z)\\
&\quad= p^{-k(2n-i-j+1)+(n-j)(n-i+1)} \sum_{M\in\Gamma(\delta_{i,j})\backslash\Gamma_n}\\
&\qquad\times\sum_{\lambda\in L_{i,j}} 1|_{k,\mathcal{M}}([(\lambda,0),0_2],M)(\tau,z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)) \sum_{\substack{x={}^tx\in(\mathbb{Z}/p\mathbb{Z})^{(n,n)}\\ x=diag(0_i,x',0_{n-j})\\ rank_p(x')=\beta}} e\left(\frac{1}{p}\mathcal{M}^t\lambda x\lambda\right),
\end{aligned}$$

where x runs over a complete set of representatives of $(\mathbb{Z}/p\mathbb{Z})^{(n,n)}$ such that $x={}^tx$, $rank_p(x)=\beta$ and $x=diag(0_i,x',0_{n-j})$ with some $x'\in(\mathbb{Z}/p\mathbb{Z})^{(j-i,j-i)}$.

Proof. We proceed as in [Yk 89, Proposition 3.2]. The inside of the last summation of the definition of $K_{i,j}^{\beta}(\tau,z)$ is

$$
\begin{aligned}
&\left(1|_{k,\mathcal{M}}([(\lambda,0),0_2],\delta_{i,j}(x)M\times\left(\begin{smallmatrix}U&0\\0&p^2U^{-1}\end{smallmatrix}\right))\right)(\tau,z)\\
&=\det(p^2U^{-1})^{-k}\det(\delta_{i,j})^{-k}\\
&\quad\times\left(e(\mathcal{M}({}^t\lambda(p^2\delta_{ij}^{-1}\tau+x)\delta_{ij}^{-1}\lambda+2{}^t\lambda\delta_{ij}^{-1}z\begin{pmatrix}p^2&\\&p\end{pmatrix}))|_{k,\mathcal{M}[\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)]}M\right)(\tau,z)\\
&=p^{-k(2n-i-j+1)}\\
&\times\left(\left((1|_{k,\mathcal{M}}([(p\delta_{i,j}^{-1}\lambda,0),0_2],\begin{pmatrix}1&p^{-1}x\\0&1\end{pmatrix}))(\tau,z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right))\right)|_{k,\mathcal{M}[\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)]}M\right)(\tau,z)\\
&=p^{-k(2n-i-j+1)}\left(1|_{k,\mathcal{M}}([(p\delta_{i,j}^{-1}\lambda,0),0_2],\begin{pmatrix}1&p^{-1}x\\0&1\end{pmatrix}M)\right)(\tau,z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)).
\end{aligned}
$$

Here we used the identity $\delta_{i,j}x=\delta_{i,j}\mathrm{diag}(0_i,x',0_{n-j})=px$. Thus

$$
\begin{aligned}
K_{i,j}^{\beta}(\tau,z) &= p^{-k(2n-i-j+1)}\sum_{\substack{[x]\\ rank_p(x)=\beta}}\ \sum_{M\in\Gamma(\delta_{i,j}(x))\backslash\Gamma_n}\\
&\quad\times\sum_{\lambda\in\mathbb{Z}^n}1|_{k,\mathcal{M}}\left([(p\delta_{i,j}^{-1}\lambda,0),0_2],\begin{pmatrix}1&p^{-1}x\\0&1\end{pmatrix}M\right)(\tau,z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)).
\end{aligned}
$$

We put

$$
\mathcal{U}\ :=\ \left\{\begin{pmatrix}1_n&s\\0_n&1_n\end{pmatrix}\middle|\ s={}^ts\in\mathbb{Z}^{(n,n)}\right\}.
$$

Then the set

$$
\mathcal{V}\ :=\ \left\{\begin{pmatrix}1_n&s\\0_n&1_n\end{pmatrix}\middle|\ {\scriptstyle s=\left(\begin{smallmatrix}0&0&0\\0&0&s_2\\0&{}^ts_2&s_3\end{smallmatrix}\right),\ s_2\in(\mathbb{Z}/p\mathbb{Z})^{(j-i,n-j)},\ s_3={}^ts_3\in(\mathbb{Z}/p\mathbb{Z})^{(n-j,n-j)}}\right\}
$$

is a complete set of representatives of $\Gamma(\delta_{i,j}(x))\backslash\Gamma(\delta_{i,j}(x))\mathcal{U}$. Therefore

$$
\begin{aligned}
&K_{i,j}^{\beta}(\tau,z)\\
&=p^{-k(2n-i-j+1)}\sum_{\substack{[x]\\ rank_p(x)=\beta}}\ \sum_{M\in(\Gamma(\delta_{i,j}(x))\mathcal{U})\backslash\Gamma_n}\ \sum_{\lambda\in\mathbb{Z}^{(n,2)}}\ \sum_{\left(\begin{smallmatrix}1_n&s\\0&1_n\end{smallmatrix}\right)\in\mathcal{V}}\\
&\quad\times\quad 1|_{k,\mathcal{M}}([(p\delta_{i,j}^{-1}\lambda,0),0_2],\begin{pmatrix}1_n&p^{-1}x\\0&1_n\end{pmatrix}\begin{pmatrix}1_n&s\\0&1_n\end{pmatrix}M)(\tau,z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right))
\end{aligned}
$$

Hence

$$K^{\beta}_{i,j}(\tau,z) = p^{-k(2n-i-j+1)} \sum_{\substack{[x] \\ rank_p(x)=\beta}} \sum_{M\in(\Gamma(\delta_{i,j}(x))\mathcal{U})\backslash\Gamma_n} \sum_{\lambda\in\mathbb{Z}^{(n,2)}}$$
$$\times\ 1|_{k,\mathcal{M}}([(p\delta_{i,j}^{-1}\lambda,0),0_2],\begin{pmatrix}1_n & p^{-1}x\\ 0 & 1_n\end{pmatrix}M)(\tau,z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right))$$
$$\times \sum_{\left(\begin{smallmatrix}1_n & s\\ 0 & 1_n\end{smallmatrix}\right)\in\mathcal{V}} e\left(p^2\mathcal{M}{}^t\lambda\delta_{i,j}^{-1}s\delta_{i,j}^{-1}\lambda\right).$$

The last summation of the RHS of the above identity is

$$\sum_{\left(\begin{smallmatrix}1_n & s\\ 0 & 1_n\end{smallmatrix}\right)\in\mathcal{V}} e\left(p^2\mathcal{M}{}^t\lambda\delta_{i,j}^{-1}s\delta_{i,j}^{-1}\lambda\right)$$
$$= \begin{cases} p^{(n-j)(n-i+1)} & \text{if } \lambda_3\mathcal{M}{}^t\lambda_3\equiv 0 \mod p^2 \text{ and } 2\lambda_3\mathcal{M}{}^t\lambda_2\equiv 0 \mod p,\\ 0 & \text{otherwise,}\end{cases}$$

where $\lambda = \left(\begin{smallmatrix}\lambda_1\\ \lambda_2\\ \lambda_3\end{smallmatrix}\right)\in\mathbb{Z}^{(n,2)}$ with $\lambda_1\in\mathbb{Z}^{(i,2)}$, $\lambda_2\in\mathbb{Z}^{(j-i,2)}$ and $\lambda_3\in\mathbb{Z}^{(n-j,2)}$. Thus

$$K^{\beta}_{i,j}(\tau,z) = p^{-k(2n-i-j+1)+(n-j)(n-i+1)} \sum_{\substack{[x] \\ rank_p(x)=\beta}} \sum_{M\in(\Gamma(\delta_{i,j}(x))\mathcal{U})\backslash\Gamma_n}$$
$$\times \sum_{\lambda\in L_{i,j}} 1|_{k,\mathcal{M}}([(\lambda,0),0_2],\left(\begin{smallmatrix}1_n & p^{-1}x\\ 0 & 1_n\end{smallmatrix}\right)M)(\tau,z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)).$$

Now $\Gamma(\delta_{i,j}(x))\mathcal{U}$ is a subgroup of $\Gamma(\delta_{i,j})$. For any $\left(\begin{smallmatrix}A & B\\ 0_n & {}^tA^{-1}\end{smallmatrix}\right)\in\Gamma(\delta_{i,j})$ we have

$$\begin{aligned}
&1|_{k,\mathcal{M}}([(\lambda,0),0_2],\left(\begin{smallmatrix}1_n & p^{-1}x\\ 0_n & 1_n\end{smallmatrix}\right)\left(\begin{smallmatrix}A & B\\ 0_n & {}^tA^{-1}\end{smallmatrix}\right)M)\\
&= 1|_{k,\mathcal{M}}([(\lambda,0),0_2],\left(\begin{smallmatrix}A & B\\ 0_n & {}^tA^{-1}\end{smallmatrix}\right)\left(\begin{smallmatrix}1_n & p^{-1}A^{-1}x{}^tA^{-1}\\ 0_n & 1_n\end{smallmatrix}\right)M)\\
&= 1|_{k,\mathcal{M}}([({}^tA\lambda,{}^tB\lambda),0_2],\left(\begin{smallmatrix}1_n & p^{-1}A^{-1}x{}^tA^{-1}\\ 0_n & 1_n\end{smallmatrix}\right)M)\\
&= 1|_{k,\mathcal{M}}([({}^tA\lambda,0),0_2],\left(\begin{smallmatrix}1_n & p^{-1}A^{-1}x{}^tA^{-1}\\ 0_n & 1_n\end{smallmatrix}\right)M),
\end{aligned}$$

and ${}^tAL_{i,j} = L_{i,j}$. Moreover, when $\left(\begin{smallmatrix}A & B\\ 0_n & {}^tA^{-1}\end{smallmatrix}\right)$ runs over all elements in a complete set of representatives of $\Gamma(\delta_{i,j}(x))\mathcal{U}\backslash\Gamma(\delta_{i,j})$, then $A^{-1}x{}^tA^{-1}$ runs over all elements in the equivalence class $[x]$ (cf. [Yk 89, proof of Proposition 3.2]).

Therefore we have

$$\begin{aligned}
&K^{\beta}_{i,j}(\tau,z)\\
&= p^{-k(2n-i-j+1)+(n-j)(n-i+1)} \sum_{\substack{x={}^t x\in(\mathbb{Z}/p\mathbb{Z})^{(n,n)}\\ x=diag(0_i,x',0_{n-j})\\ rank_p(x')=\beta}} \sum_{M\in\Gamma(\delta_{i,j})\backslash\Gamma_n}\\
&\qquad \times \sum_{\lambda\in L_{i,j}} 1|_{k,\mathcal{M}}([(\lambda,0),0_2],\left(\begin{smallmatrix}1_n & p^{-1}x\\ 0 & 1_n\end{smallmatrix}\right)M)(\tau,z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right))\\
&= p^{-k(2n-i-j+1)+(n-j)(n-i+1)} \sum_{M\in\Gamma(\delta_{i,j})\backslash\Gamma_n}\\
&\qquad \times \sum_{\lambda\in L_{i,j}} 1|_{k,\mathcal{M}}([(\lambda,0),0_2],M)(\tau,z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)) \sum_{\substack{x={}^t x\in(\mathbb{Z}/p\mathbb{Z})^{(n,n)}\\ x=diag(0_i,x',0_{n-j})\\ rank_p(x')=\beta}} e\left(\frac{1}{p}\mathcal{M}{}^t\lambda x\lambda\right).
\end{aligned}$$

$\square$

5.2 The function $\tilde{K}^{\beta}_{i,j}$

The purpose of this subsection is to introduce a function $\tilde{K}^{\beta}_{i,j}$ and to express $E^{(n)}_{k,\mathcal{M}}|V_{\alpha,n-\alpha}(p^2)$ as a summation of $\tilde{K}^{\beta}_{i,j}$. Moreover, we shall show that $\tilde{K}^{\beta}_{i,j}$ is a summation of exponential functions with generalized Gauss sums (cf. Proposition 5.3).

We define

$$\begin{aligned}
L^*_{i,j} &:= L^*_{i,j,\mathcal{M},p}\\
&= \left\{ \left(\begin{smallmatrix}\lambda_1\\ \lambda_2\\ \lambda_3\end{smallmatrix}\right) \in L_{i,j} \,\middle|\, 2\lambda_3\mathcal{M}\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)\in\mathbb{Z}^{(n-j,1)} \right\}\\
&= \left\{ \begin{pmatrix}\lambda_1\\ \lambda_2\\ \lambda_3\end{pmatrix} \in (p^{-1}\mathbb{Z})^{(n,2)} \,\middle|\, \begin{array}{c} \lambda_1\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)^{-1}\in\mathbb{Z}^{(i,2)},\ \lambda_2\in\mathbb{Z}^{(j-i,2)}\\ \lambda_3\in(p^{-1}\mathbb{Z})^{(n-j,2)},\ 2\lambda_2\mathcal{M}{}^t\lambda_3\in\mathbb{Z}^{(j-i,n-j)}\\ \lambda_3\mathcal{M}{}^t\lambda_3\in\mathbb{Z}^{(n-j,n-j)},\ 2\lambda_3\mathcal{M}\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)\in\mathbb{Z}^{(n-j,1)}\end{array} \right\}
\end{aligned}$$

and define a generalized Gauss sum

$$G^{j-i,l}_{\mathcal{M}}(\lambda_2) \quad := \sum_{\substack{x'={}^t x'\in(\mathbb{Z}/p\mathbb{Z})^{(j-i,j-i)}\\ rank_p(x')=j-i-l}} e\left(\frac{1}{p}\mathcal{M}{}^t\lambda_2 x'\lambda_2\right)$$

for $\lambda_2\in\mathbb{Z}^{(j-i,2)}$. We define

$$\begin{aligned}
&\tilde{K}^{\beta}_{i,j}(\tau,z) := \tilde{K}^{\beta}_{i,j,\mathcal{M},p}(\tau,z)\\
&\qquad\qquad = \sum_{u,v\in(\mathbb{Z}/p\mathbb{Z})^{(n,1)}} \left(K^{\beta}_{i,j}|_{k,\mathcal{M}[\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)]}[((0,u),(0,v)),0_2]\right)(\tau,z).
\end{aligned}$$

Proposition 5.3. *Let the notation be as above. Then we obtain*

$$\left(E^{(n)}_{k,\mathcal{M}}|V_{\alpha,n-\alpha}(p^2)\right)(\tau,z) \quad = \sum_{\substack{i,j\\ 0\le i\le j\le n\\ j-i\ge n-\alpha}} \tilde{K}^{\alpha-i-n+j}_{i,j}(\tau,z),$$

where

$$\begin{aligned}\tilde{K}^{\alpha-i-n+j}_{i,j}(\tau,z) &= p^{-k(2n-i-j+1)+(n-j)(n-i+1)+2n-j}\\ &\times \sum_{M\in\Gamma(\delta_{i,j})\backslash\Gamma_n} \sum_{\lambda=\begin{pmatrix}\lambda_1\\ \lambda_2\\ \lambda_3\end{pmatrix}\in L^*_{i,j}} \{1|_{k,\mathcal{M}}([(\lambda,0),0_2],M)\}\left(\tau,z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)\right)\\ &\quad\times \sum_{u_2\in(\mathbb{Z}/p\mathbb{Z})^{(j-i,1)}} G^{j-i,n-\alpha}_{\mathcal{M}}(\lambda_2+(0,u_2)).\end{aligned}$$

Proof. From the definition of $\tilde{K}^{\beta}_{i,j}$ and Lemma 5.2 we obtain

$$\begin{aligned}&\tilde{K}^{\alpha-i-n+j}_{i,j}(\tau,z)\\ &= p^{-k(2n-i-j+1)+(n-j)(n-i+1)} \sum_{M\in\Gamma(\delta_{i,j})\backslash\Gamma_n} \sum_{\lambda=\begin{pmatrix}\lambda_1\\ \lambda_2\\ \lambda_3\end{pmatrix}\in L_{i,j}} G^{j-i,n-\alpha}_{\mathcal{M}}(\lambda_2)\\ &\times \sum_{u,v\in(\mathbb{Z}/p\mathbb{Z})^{(n,1)}} \left(1|_{k,\mathcal{M}}([(\lambda,0),0_2],M)\left(\tau,z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)\right)\right)|_{k,\mathcal{M}\left[\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)\right]}[((0,u),(0,v)),0_2],\end{aligned} \tag{5.1}$$

where $\lambda_1 \in \mathbb{Z}^{(i,2)}$, $\lambda_2 \in \mathbb{Z}^{(j-i,2)}$ and $\lambda_3 \in \mathbb{Z}^{(n-j,2)}$ satisfy $\begin{pmatrix}\lambda_1\\ \lambda_2\\ \lambda_3\end{pmatrix} \in L_{i,j}$, and where the $n\times 2$ matrix $\lambda = \begin{pmatrix}\lambda_1\\ \lambda_2\\ \lambda_3\end{pmatrix}$ runs over the set $L_{i,j}$.
By a straightforward calculation we have

$$\begin{aligned}&(1|_{k,\mathcal{M}}([(\lambda,0),0_2],M))\left(\tau,z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)\right)\\ &= \left(1|_{k,\mathcal{M}\left[\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)\right]}\left(\left[\left(\lambda\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)^{-1},0\right),0_2\right],M\right)\right)(\tau,z).\end{aligned}$$

Thus the last summation of (5.1) is

$$\begin{aligned}&\sum_{u,v\in(\mathbb{Z}/p\mathbb{Z})^{(n,1)}} \left\{1|_{k,\mathcal{M}}([(\lambda,0),0_2],M)\left(\tau,z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)\right)|_{k,\mathcal{M}\left[\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)\right]}[((0,u),(0,v)),0_2]\right\}(\tau,z)\\ &= \sum_{u,v\in(\mathbb{Z}/p\mathbb{Z})^{(n,1)}}\\ &\quad\times\left\{1|_{k,\mathcal{M}\left[\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)\right]}\left(\left[\left(\lambda\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)^{-1},0\right),0_2\right],M\right)|_{k,\mathcal{M}\left[\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)\right]}[((0,u),(0,v)),0_2]\right\}(\tau,z)\end{aligned}$$

$$= \sum_{u',v' \in (\mathbb{Z}/p\mathbb{Z})^{(n,1)}} \left\{ 1|_{k,\mathcal{M}[\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)]}([(\lambda \left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)^{-1} + (0,u'),(0,v')),0_2],M) \right\} (\tau, z)$$

$$= \sum_{u',v' \in (\mathbb{Z}/p\mathbb{Z})^{(n,1)}} \left\{ 1|_{k,\mathcal{M}}([(\lambda + (0,u'),(0,v')),0_2],M) \right\} (\tau, z \left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right))$$

$$= \sum_{u' \in (\mathbb{Z}/p\mathbb{Z})^{(n,1)}} \left\{ 1|_{k,\mathcal{M}}([(\lambda + (0,u'),0),0_2],M) \right\} (\tau, z \left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)) \sum_{v' \in (\mathbb{Z}/p\mathbb{Z})^{(n,1)}} e(2\mathcal{M}\,{}^t\lambda(0,v')),$$

where, in the second identity, we used

$$(M, [((0,u),(0,v)),0_2]) \quad = \quad ([((0,u'),(0,v')),0_2],M)$$

with $\left(\begin{smallmatrix} u' \\ v' \end{smallmatrix}\right) = \left(\begin{smallmatrix} D & -C \\ -B & A \end{smallmatrix}\right) \left(\begin{smallmatrix} u \\ v \end{smallmatrix}\right)$ for $M = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right) \in \Gamma_n$. For $\lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} \in L_{i,j}$ we now have

$$\sum_{v' \in (\mathbb{Z}/p\mathbb{Z})^{(n,1)}} e(2\mathcal{M}\,{}^t\lambda(0,v')) \quad = \quad \begin{cases} p^n & \text{if } 2\lambda_3\mathcal{M}\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) \in \mathbb{Z}^{(n-j,1)}, \\ 0 & \text{otherwise.} \end{cases}$$

Therefore

$$\tilde{K}_{i,j}^{\alpha-i-n+j}(\tau,z)$$
$$= p^{-k(2n-i-j+1)+(n-j)(n-i+1)+n} \sum_{M \in \Gamma(\delta_{i,j}) \backslash \Gamma_n} \sum_{\substack{\lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} \in L_{i,j} \\ 2\lambda_3\mathcal{M}\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) \in \mathbb{Z}^{(n-j,1)}}} G_{\mathcal{M}}^{j-i,n-\alpha}(\lambda_2)$$
$$\times \sum_{u \in (\mathbb{Z}/p\mathbb{Z})^{(n,1)}} \left\{ 1|_{k,\mathcal{M}}([(\lambda + (0,u),0),0_2],M) \right\} (\tau, z \left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)).$$

Thus

$$\tilde{K}_{i,j}^{\alpha-i-n+j}(\tau,z)$$
$$= p^{-k(2n-i-j+1)+(n-j)(n-i+1)+n} \sum_{M \in \Gamma(\delta_{i,j}) \backslash \Gamma_n} \sum_{\lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} \in L_{i,j}^*}$$
$$\times \left\{ 1|_{k,\mathcal{M}}([(\lambda,0),0_2],M) \right\} (\tau, z \left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right))$$
$$\times p^{n-j} \sum_{u_2 \in (\mathbb{Z}/p\mathbb{Z})^{(j-i,1)}} G_{\mathcal{M}}^{j-i,n-\alpha}(\lambda_2 + (0,u_2)),$$

where $L_{i,j}^*$ is defined as before. □

We put

$$g_p(n,\alpha) \quad := \quad \prod_{j=1}^{\alpha} \left\{ (p^{n-j+1}-1)(p^j-1)^{-1} \right\}$$

It is not difficult to see $g_p(n, n-\alpha) = g_p(n,\alpha)$.

Lemma 5.4. *For any $\lambda = (\lambda_1, \lambda_2) \in \mathbb{Z}^{(n,2)}$ and for any prime p, we have*

$$\sum_{u_2 \in (\mathbb{Z}/p\mathbb{Z})^{(n,1)}} G^{n,\alpha}_{\mathcal{M}}(\lambda + (0, u_2))$$
$$= \begin{cases} p^{\frac{1}{4}(n-\alpha-1)^2+\frac{1}{2}(n-\alpha-1)+\alpha+n}\left(\frac{-m}{p}\right) & \\ \quad \times g_p(n-1,\alpha) \prod\limits_{\substack{j=1 \\ j:odd}}^{n-\alpha-2} (p^j - 1) & \begin{array}{l} \text{if } n-\alpha \equiv 1 \mod 2 \\ \text{and } \lambda_1 \not\equiv 0 \mod p, \end{array} \\ 0 & \begin{array}{l} \text{if } n-\alpha \equiv 1 \mod 2 \\ \text{and } \lambda_1 \equiv 0 \mod p, \end{array} \\ p^{\frac{1}{4}(n-\alpha)^2+\frac{1}{2}(n-\alpha)+\alpha} g_p(n,\alpha) \prod\limits_{\substack{j=1 \\ j:odd}}^{n-\alpha-1} (p^j - 1) & \text{if } n-\alpha \equiv 0 \mod 2. \end{cases}$$

Here $m = \det(2\mathcal{M})$ and we regard the product $\prod\limits_{\substack{j=1 \\ j:odd}}^{c} (p^j - 1)$ as 1, if c is less than 1.

Proof. This calculation is similar to the calculation of

$$\sum_{\substack{x={}^t x \in (\mathbb{Z}/p\mathbb{Z})^n \\ rank_p x = n-\alpha}} e\left(\frac{1}{p} m\, {}^t\lambda_1 x \lambda_1\right)$$

for $\lambda_1 \in \mathbb{Z}^{(n,1)}$ and for $m \in \mathbb{Z}$ which is in [Yk 89, Lemma 3.1].
If p is an odd prime and if $\lambda_1 \not\equiv 0 \mod p$, then

$$\sum_{u_2 \in (\mathbb{Z}/p\mathbb{Z})^{(n,1)}} G^{n,\alpha}_{\mathcal{M}}(\lambda + (0, u_2))$$
$$= \sum_{u_2 \in (\mathbb{Z}/p\mathbb{Z})^{(n,1)}} \sum_{\substack{x'={}^t x' \in (\mathbb{Z}/p\mathbb{Z})^{(n,n)} \\ rank_p(x') = n-\alpha}} e\left(\frac{1}{p} \mathcal{M}\, {}^t(\lambda_1, u_2) x' (\lambda_1, u_2)\right)$$

By diagonalizing the matrices x' we have

$$\sum_{u_2 \in (\mathbb{Z}/p\mathbb{Z})^{(n,1)}} G^{n,\alpha}_{\mathcal{M}}(\lambda + (0, u_2))$$
$$= \sum_{i=0,1} p^{n-1} \left|\mathrm{GL}_{n-1}(\mathbb{Z}/p\mathbb{Z})\right| \left|O(x_i)\right|^{-1}$$
$$\times \sum_{u_2 \in (\mathbb{Z}/p\mathbb{Z})^{(n,1)}} \sum_{\substack{\eta \in (\mathbb{Z}/p\mathbb{Z})^{(n,1)} \\ \eta \not\equiv 0 \mod p}} e\left(\frac{1}{p} \mathcal{M}\, {}^t(\eta, u_2) x_i (\eta, u_2)\right),$$

where $x_i = \begin{pmatrix} y_i & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{Z}^{(n,n)}$, $y_0 = 1_{n-\alpha}$, $y_1 = \begin{pmatrix} 1_{n-\alpha-1} & 0 \\ 0 & \gamma \end{pmatrix} \in \mathbb{Z}^{(n-\alpha,n-\alpha)}$ and γ is an integer such that $\left(\frac{\gamma}{p}\right) = -1$. Here $O(x_i)$ is the orthogonal group of x_i:

$$O(x_i) \quad := \quad \left\{ g \in \mathrm{GL}_n(\mathbb{Z}/p\mathbb{Z}) \,|\, g x_i {}^t g = x_i \right\}.$$

If we diagonalize the matrix $\mathcal{M}$ as $\mathcal{M} \equiv {}^t X \begin{pmatrix} m & 0 \\ 0 & 1 \end{pmatrix} X \mod p$ with $X = \begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix}$, then

$$\begin{aligned} &\sum_{u_2 \in (\mathbb{Z}/p\mathbb{Z})^{(n,1)}} G_{\mathcal{M}}^{n,\alpha}(\lambda + (0,u_2)) \\ &= \sum_{i=0,1} p^{n-1} \left|\mathrm{GL}_{n-1}(\mathbb{Z}/p\mathbb{Z})\right| \left|O(x_i)\right|^{-1} \\ &\quad \times \sum_{u_2 \in (\mathbb{Z}/p\mathbb{Z})^{(n,1)}} \sum_{\substack{\eta \in (\mathbb{Z}/p\mathbb{Z})^{(n,1)} \\ \eta \not\equiv 0 \mod p}} e\left(\frac{1}{p} \left(m \eta {}^t \eta + u_2 {}^t u_2 \right) x_i \right). \end{aligned}$$

The rest of the calculation is an analogue to [Yk 89, Lemma 3.1]. For the case of $p = 2$ or $\lambda_1 \equiv 0 \mod p$, the calculation is similar. If $p = 2$, we need to calculate the case that $\mathcal{M} = {}^t X \begin{pmatrix} m' & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix} X$, but it is not difficult. We leave the detail to the reader. $\square$

We set

$$S_{\mathcal{M}}^{n,\alpha}(0) := \sum_{u_2 \in (\mathbb{Z}/p\mathbb{Z})^{(n,1)}} G_{\mathcal{M}}^{n,\alpha}((0,u_2))$$

and

$$S_{\mathcal{M}}^{n,\alpha}(1) := \sum_{u_2 \in (\mathbb{Z}/p\mathbb{Z})^{(n,1)}} G_{\mathcal{M}}^{n,\alpha}\left(\left(\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, u_2 \right) \right).$$

Due to Lemma 5.4, we have that $\displaystyle\sum_{u_2 \in (\mathbb{Z}/p\mathbb{Z})^{(n,1)}} G_{\mathcal{M}}^{n,\alpha}(\lambda + (0,u_2))$ equals $S_{\mathcal{M}}^{n,\alpha}(0)$ or $S_{\mathcal{M}}^{n,\alpha}(1)$, according as $\lambda \in \mathbb{Z}^{(n,2)} \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}$ or $\lambda \notin \mathbb{Z}^{(n,2)} \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}$.

Proposition 5.5. *The form $E_{k,\mathcal{M}}^{(n)} | V_{\alpha,n-\alpha}(p^2)$ is a linear combination of three forms $E^{(n)}_{k,\mathcal{M}\left[\begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}\right]}$, $E_{k,\mathcal{M}}^{(n)} | U_{\begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}}$ and $E^{(n)}_{k,\mathcal{M}\left[X^{-1}\begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}^{-1}\right]} | U_{\begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix} X \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}}$. Here the index-shift map U_L is defined in §4.6, and $X = \begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix}$ is a matrix in $\mathbb{Z}^{(2,2)}$ such that $\mathcal{M} = {}^t X \begin{pmatrix} m+1 & 1 \\ 1 & 1 \end{pmatrix} X$ if $p = 2$ and $\frac{\det(2\mathcal{M})}{4} \equiv 3 \mod 4$, or $\mathcal{M} \equiv {}^t X \begin{pmatrix} m & 0 \\ 0 & 1 \end{pmatrix} X \mod p$ otherwise, and where $m = \det(2\mathcal{M})$.*

Proof. By virtue of Proposition 5.3 we only need to show that the form $\tilde{K}_{i,j}^{\alpha-i-n+j}(\tau,z)$ is a linear combination of the above three forms.

Because of the conditions $\lambda_3\mathcal{M}{}^t\lambda_3 \in \mathbb{Z}^{(n-j,n-j)}$ and $2\lambda_3\mathcal{M}\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right) \in \mathbb{Z}^{(n-j,1)}$ in the definition of $L^*_{i,j}$, we obtain

$$L^*_{i,j} = \left\{ \begin{pmatrix}\lambda_1\\\lambda_2\\\lambda_3\end{pmatrix} \in \left(\frac{1}{p}\mathbb{Z}\right)^{(n,2)} \middle| \begin{array}{c} \lambda_1 \in \mathbb{Z}^{(i,2)}\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right), \lambda_2 \in \mathbb{Z}^{(j-i,2)}, \\ \lambda_3{}^tX \in \mathbb{Z}^{(n-j,2)}\begin{pmatrix}p^{-1}&0\\0&1\end{pmatrix}\end{array} \right\} \quad (5.2)$$

for the case $p|f$, and

$$L^*_{i,j} = \left\{ \begin{pmatrix}\lambda_1\\\lambda_2\\\lambda_3\end{pmatrix} \in \mathbb{Z}^{(n,2)} \,|\, \lambda_1 \in \mathbb{Z}^{(i,2)}\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right), \lambda_2 \in \mathbb{Z}^{(j-i,2)}, \lambda_3 \in \mathbb{Z}^{(n-j,2)} \right\} \quad (5.3)$$

for the case $p\not| f$. Here f is a natural number such that $D_0f^2 = -\det(2\mathcal{M})$ and D_0 is a fundamental discriminant, and where the matrix X is stated in this proposition.

We now assume $p|f$. If p is an odd prime, then the matrix $X = \left(\begin{smallmatrix}1&0\\x&1\end{smallmatrix}\right) \in \mathbb{Z}^{(2,2)}$ satisfies $\mathcal{M} \equiv {}^tX\left(\begin{smallmatrix}m&0\\0&1\end{smallmatrix}\right)X \mod p$ and $p^2|m$. If $p=2$, then the matrix $X = \left(\begin{smallmatrix}1&0\\x&1\end{smallmatrix}\right) \in \mathbb{Z}^{(2,2)}$ satisfies $\mathcal{M} = {}^tX\left(\begin{smallmatrix}m&0\\0&1\end{smallmatrix}\right)X$ with $4|m$, or $\mathcal{M} = {}^tX\left(\begin{smallmatrix}m'&1\\1&1\end{smallmatrix}\right)X$ with $4|m'$. We remark that $\mathcal{M}\left[X^{-1}\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)^{-1}\right]$ is a half-integral symmetric matrix.

We put

$$\begin{aligned} L_0 &:= \left\{ \begin{pmatrix}\lambda_1\\\lambda_2\\\lambda_3\end{pmatrix} \in L^*_{i,j} \,|\, \lambda_2 \in \mathbb{Z}^{(j-i,2)}\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right) \right\}, \\ L_1 &:= \left\{ \begin{pmatrix}\lambda_1\\\lambda_2\\\lambda_3\end{pmatrix} \in L^*_{i,j} \,|\, \lambda_2 \not\in \mathbb{Z}^{(j-i,2)}\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right) \right\}. \end{aligned}$$

and set

$$L'_{i,j} := \left\{ \begin{pmatrix}\lambda_1\\\lambda_2\\\lambda_3\end{pmatrix} \middle| \lambda_1 \in \mathbb{Z}^{(i,2)}\begin{pmatrix}p^2&0\\0&1\end{pmatrix}, \lambda_2 \in \mathbb{Z}^{(j-i,2)}\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right), \lambda_3 \in \mathbb{Z}^{(n-j,2)} \right\}.$$

By using the identity

$$\begin{aligned} &\{1|_{k,\mathcal{M}}([(\lambda,0),0_2],M)\}(\tau, z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)) \\ &= \left\{1\,\middle|_{k,\mathcal{M}[X^{-1}]}([(\lambda{}^tX,0),0_2],M)\right\}(\tau, z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right){}^tX) \\ &= \left\{1\,\middle|_{k,\mathcal{M}\left[X^{-1}\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)^{-1}\right]}([(\lambda{}^tX\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right),0),0_2],M)\right\}(\tau, z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right){}^tX\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)), \end{aligned}$$

we have

$$
\begin{aligned}
&\tilde{K}_{i,j}^{\alpha-i-n+j}(\tau,z)\\
=\ & p^{-k(2n-i-j+1)+(n-j)(n-i+1)+2n-j}\\
&\times \sum_{M\in\Gamma(\delta_{i,j})\backslash\Gamma_n}\Bigg\{S_{\mathcal{M}}^{j-i,n-\alpha}(0)\sum_{\lambda\in L_0}\{1|_{k,\mathcal{M}}([(\lambda,0),0_2],M)\}\left(\tau,z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)\right)\\
&+S_{\mathcal{M}}^{j-i,n-\alpha}(1)\sum_{\lambda\in L_1}\{1|_{k,\mathcal{M}}([(\lambda,0),0_2],M)\}\left(\tau,z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)\right)\Bigg\}\\
=\ & p^{-k(2n-i-j+1)+(n-j)(n-i+1)+2n-j}\sum_{M\in\Gamma(\delta_{i,j})\backslash\Gamma_n}\Bigg\{\left(S_{\mathcal{M}}^{j-i,n-\alpha}(0)-S_{\mathcal{M}}^{j-i,n-\alpha}(1)\right)\\
&\times\sum_{\lambda\in L_0}\left\{1\Big|_{k,\mathcal{M}\left[X^{-1}\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)^{-1}\right]}\left(\left[\left(\lambda^tX\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right),0\right),0_2\right],M\right)\right\}\left(\tau,z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right){}^tX\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)\right)\\
&+S_{\mathcal{M}}^{j-i,n-\alpha}(1)\\
&\times\sum_{\lambda\in L_{i,j}^*}\left\{1\Big|_{k,\mathcal{M}\left[X^{-1}\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)^{-1}\right]}\left(\left[\left(\lambda^tX\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right),0\right),0_2\right],M\right)\right\}\left(\tau,z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right){}^tX\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)\right)\Bigg\}\\
=\ & p^{-k(2n-i-j+1)+(n-j)(n-i+1)+2n-j}\sum_{M\in\Gamma(\delta_{i,j})\backslash\Gamma_n}\Bigg\{\left(S_{\mathcal{M}}^{j-i,n-\alpha}(0)-S_{\mathcal{M}}^{j-i,n-\alpha}(1)\right)\\
&\times\sum_{\lambda\in L'_{j,j}}\left\{1\Big|_{k,\mathcal{M}\left[X^{-1}\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)^{-1}\right]}([(\lambda,0),0_2],M)\right\}\left(\tau,z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right){}^tX\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)\right)\\
&+S_{\mathcal{M}}^{j-i,n-\alpha}(1)\\
&\times\sum_{\lambda\in L'_{i,j}}\left\{1\Big|_{k,\mathcal{M}\left[X^{-1}\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)^{-1}\right]}([(\lambda,0),0_2],M)\right\}\left(\tau,z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right){}^tX\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)\right)\Bigg\}.
\end{aligned}
$$

We now calculate the sum

$$
\sum_{M\in\Gamma(\delta_{i,j})\backslash\Gamma_n}\sum_{\lambda\in L'_{j,j}}\left\{1\Big|_{k,\mathcal{M}\left[X^{-1}\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)^{-1}\right]}([(\lambda,0),0_2],M)\right\}\left(\tau,z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right){}^tX\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)\right).
$$

We set

$$
H_{i,j}\quad := \delta_{i,j}\,\mathrm{GL}_n(\mathbb{Z})\,\delta_{i,j}^{-1}\cap\mathrm{GL}_n(\mathbb{Z}).
$$

If $\{A_l\}_l$ is a complete set of representatives of $H_{i,j}\backslash\mathrm{GL}_n(\mathbb{Z})$, then one can say that the set $\left\{\begin{pmatrix}A_l&0\\0&{}^tA_l{}^{-1}\end{pmatrix}\right\}_l$ is a complete set of representatives of

$\Gamma(\delta_{i,j})\backslash\Gamma_\infty^{(n)}$. Thus

$$\begin{aligned}
&\sum_{M\in\Gamma(\delta_{i,j})\backslash\Gamma_n}\sum_{\lambda\in L'_{j,j}}\{1|_{k,\mathcal{M}}([(\lambda,0),0_2],M)\}(\tau,z)\\
&=\sum_{M\in\Gamma_\infty^{(n)}\backslash\Gamma_n}\sum_{A\in H_{i,j}\backslash GL_n(\mathbb{Z})}\sum_{\lambda\in L'_{j,j}}\left\{1|_{k,\mathcal{M}}([(\lambda,0),0_2],\left(\begin{smallmatrix}A&0\\0&{}^tA^{-1}\end{smallmatrix}\right)M)\right\}(\tau,z)\\
&=\sum_{M\in\Gamma_\infty^{(n)}\backslash\Gamma_n}\sum_{A\in H_{i,j}\backslash GL_n(\mathbb{Z})}\sum_{\lambda\in L'_{j,j}}\left\{1|_{k,\mathcal{M}}([({}^tA\lambda,0),0_2],M)\right\}(\tau,z).
\end{aligned}$$

If $B(\lambda)$ is a function on $\lambda\in\mathbb{Z}^{(n,2)}$. Then

$$\begin{aligned}
&\sum_{A\in H_{i,j}\backslash GL_n(\mathbb{Z})}\sum_{\lambda\in L'_{j,j}}B({}^tA\lambda)\\
&=[H_{j,j}:H_{i,j}]\sum_{A\in H_{j,j}\backslash GL_n(\mathbb{Z})}\sum_{\lambda\in L'_{j,j}}B({}^tA\lambda)\\
&=[H_{j,j}:H_{i,j}]\left(a_0\sum_{\lambda\in\mathbb{Z}^{(n,2)}}B(\lambda)+a_1\sum_{\lambda\in\mathbb{Z}^{(n,2)}}B(\lambda\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right))+a_2\sum_{\lambda\in\mathbb{Z}^{(n,2)}}B(\lambda\left(\begin{smallmatrix}p^2&0\\0&1\end{smallmatrix}\right))\right)
\end{aligned}$$

with numbers a_0, a_1 and a_2 under the assumption that the summations converges absolutely. The values a_0, a_1 and a_2 are independent of the choice of the function B. For the exact values of a_0, of a_1 and of a_2 the reader is referred to [H 13, Lemma 3.7].
Hence we have

$$\begin{aligned}
&\sum_{M\in\Gamma(\delta_{i,j})\backslash\Gamma_n}\sum_{\lambda\in L'_{j,j}}\left\{1|_{k,\mathcal{M}\left[X^{-1}\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)^{-1}\right]}([(\lambda,0),0_2],M)\right\}(\tau,z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right){}^tX\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right))\\
=\;&[H_{j,j}:H_{i,j}]\sum_{M\in\Gamma_\infty^{(n)}\backslash\Gamma_n}\\
&\times\left(a_0\sum_{\lambda\in\mathbb{Z}^{(n,2)}}\left\{1|_{k,\mathcal{M}\left[X^{-1}\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)^{-1}\right]}([(\lambda,0),0_2],M)\right\}(\tau,z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right){}^tX\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right))\right.\\
&+a_1\sum_{\lambda\in\mathbb{Z}^{(n,2)}}\left\{1|_{k,\mathcal{M}\left[X^{-1}\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)^{-1}\right]}([(\lambda\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right),0),0_2],M)\right\}(\tau,z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right){}^tX\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right))\\
&+a_2\sum_{\lambda\in\mathbb{Z}^{(n,2)}}\\
&\left.\times\left\{1|_{k,\mathcal{M}\left[X^{-1}\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)^{-1}\right]}([(\lambda\left(\begin{smallmatrix}p^2&0\\0&1\end{smallmatrix}\right),0),0_2],M)\right\}(\tau,z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right){}^tX\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right))\right)\\
=\;&[H_{j,j}:H_{i,j}]\sum_{M\in\Gamma_\infty^{(n)}\backslash\Gamma_n}
\end{aligned}$$

$$\times\left(a_0\sum_{\lambda\in\mathbb{Z}^{(n,2)}}\left\{1|_{k,\mathcal{M}\left[X^{-1}\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)^{-1}\right]}([(\lambda,0),0_2],M)\right\}(\tau,z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right){}^tX\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right))\right.$$
$$+a_1\sum_{\lambda\in\mathbb{Z}^{(n,2)}}\{1|_{k,\mathcal{M}}([(\lambda,0),0_2],M)\}(\tau,z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right))$$
$$\left.+a_2\sum_{\lambda\in\mathbb{Z}^{(n,2)}}\left\{1|_{k,\mathcal{M}\left[\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)\right]}([(\lambda,0),0_2],M)\right\}(\tau,z)\right)$$
$$=\quad [H_{j,j}:H_{i,j}]\left(a_0E^{(n)}_{k,\mathcal{M}\left[X^{-1}\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)^{-1}\right]}(\tau,z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right){}^tX\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right))\right.$$
$$\left.+a_1E^{(n)}_{k,\mathcal{M}}(\tau,z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right))+a_2E^{(n)}_{k,\mathcal{M}\left[\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)\right]}(\tau,z)\right).$$

Similarly, the summation

$$\sum_{M\in\Gamma(\delta_{i,j})\backslash\Gamma_n}\sum_{\lambda\in L'_{i,j}}\left\{1|_{k,\mathcal{M}\left[X^{-1}\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)^{-1}\right]}([(\lambda,0),0_2],M)\right\}(\tau,z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right){}^tX\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right))$$

is a linear combination of $E^{(n)}_{k,\mathcal{M}\left[X^{-1}\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)^{-1}\right]}(\tau,z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right){}^tX\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right))$, $E^{(n)}_{k,\mathcal{M}}(\tau,z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right))$ and $E^{(n)}_{k,\mathcal{M}\left[\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)\right]}(\tau,z)$.

Therefore, if $p|f$, then the form $\tilde{K}^{\alpha-i-n+j}_{i,j}(\tau,z)$ is a linear combination of the above three forms.

The proof for the case $p\nmid f$ is similar to the case $p|f$. If $p\nmid f$, then $\tilde{K}^{\alpha-i-n+j}_{i,j}(\tau,z)$ is a linear combination of two forms $E^{(n)}_{k,\mathcal{M}}(\tau,z\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right))$ and $E^{(n)}_{k,\mathcal{M}\left[\left(\begin{smallmatrix}p&0\\0&1\end{smallmatrix}\right)\right]}(\tau,z)$. We omit the detail of the calculation here. $\square$

6 Commutativity with the Siegel operators

In [Kr 86] an explicit commutative relation between the generators of Hecke operators for Siegel modular forms and Siegel Φ-operator has been given. In this section we shall give a similar relation in the frameworks of Jacobi forms of matrix index and of Jacobi forms of half-integral weight.

Let $\mathcal{M}=\begin{pmatrix}l&\frac{r}{2}\\\frac{r}{2}&1\end{pmatrix}$ be a 2×2 matrix and put $m=\det(2\mathcal{M})$ as before.

For any Jacobi form $\phi\in J^{(n)}_{k,\mathcal{M}}$, or $\phi\in J^{(n)}_{k-\frac12,m}$ we define the Siegel Φ-operator

$$\Phi(\phi)(\tau',z')\quad:=\quad\lim_{t\to+\infty}\phi\left(\begin{pmatrix}\tau'&0\\0&\sqrt{-1}t\end{pmatrix},\begin{pmatrix}z'\\0\end{pmatrix}\right)$$

for $(\tau',z')\in\mathfrak{H}_{n-1}\times\mathbb{C}^{(n-1,2)}$, or for $(\tau',z')\in\mathfrak{H}_{n-1}\times\mathbb{C}^{(n-1,1)}$. This Siegel Φ-operator is a map from $J^{(n)}_{k,\mathcal{M}}$ to $J^{(n-1)}_{k,\mathcal{M}}$, or from $J^{(n)}_{k-\frac12,m}$ to $J^{(n-1)}_{k-\frac12,m}$, respectively.

Proposition 6.1. *For any Jacobi form $\phi \in J_{k,\mathcal{M}}^{(n)}$ and for any prime p, we have*

$$\Phi(\phi|V_{\alpha,n-\alpha}(p^2)) \;=\; \Phi(\phi)|V_{\alpha,n-\alpha}(p^2)^*,$$

where $V_{\alpha,n-\alpha}(p^2)^*$ is a map $V_{\alpha,n-\alpha}(p^2)^* : J_{k,\mathcal{M}}^{(n-1)} \to J^{(n-1)}_{k,\mathcal{M}\left[\left(\begin{smallmatrix} p & 0 \\ 0 & 1\end{smallmatrix}\right)\right]}$ given by

$$\begin{aligned} V_{\alpha,n-\alpha}(p^2)^* \;=\; & p^{\alpha+2-k}V_{\alpha,n-\alpha-1}(p^2) \\ & +p(1+p^{2n+1-2k})V_{\alpha-1,n-\alpha}(p^2) \\ & +(p^{2n-2\alpha+2}-1)p^{\alpha-k}V_{\alpha-2,n-\alpha+1}(p^2). \end{aligned}$$

Proof. We shall first show that there exists a linear combination of index-shift map $V_{\alpha,n-\alpha}(p^2)^*$ which satisfies $\Phi(\phi|V_{\alpha,n-\alpha}(p^2)) = \Phi(\phi)|V_{\alpha,n-\alpha}(p^2)^*$.
We set $U = \begin{pmatrix} p^2 & 0 \\ 0 & p\end{pmatrix}$. Let

$$\begin{aligned} \phi(\tau,z) \;&=\; \sum_{N,R} A_1(N,R)e(N\tau + R^tz), \\ (\phi|V_{\alpha,n-\alpha}(p^2))(\tau,z) \;&=\; \sum_{\hat{N},\hat{R}} A_2(\hat{N},\hat{R})e(\hat{N}\tau + \hat{R}^tz) \end{aligned}$$

be the Fourier expansions. Let $\left\{\begin{pmatrix} p^{2\,t}D_j^{\;-1} & B_{(j,l)} \\ 0_n & D_j\end{pmatrix}\right\}_{(j,l)}$ be a complete set of representatives of $\Gamma_n\backslash\Gamma_n\mathrm{diag}(1_\alpha, p1_{n-\alpha}, p^21_\alpha, p1_{n-\alpha})\Gamma_n$. Then the Fourier coefficients $A_2(\hat{N},\hat{R})$ have been calculated in the proof of Proposition 4.4:

$$A_2(\hat{N},\hat{R}) \;=\; p^{-k+n}\sum_j \det(D_j)^{-k} \sum_{\lambda_2\in(\mathbb{Z}/p\mathbb{Z})^{(n,1)}} A_1(N,R)\sum_l e(NB_{(j,l)}D_j^{-1}) \qquad (6.1)$$

Here N and R are determined by

$$\begin{aligned} N \;&=\; \frac{1}{p^2}D_j\left(\left(\hat{N}-\frac{1}{4}\hat{R}_2{}^t\hat{R}_2\right)+\frac{1}{4}(\hat{R}_2-2\lambda_2)^t(\hat{R}_2-2\lambda_2)\right){}^tD_j, \qquad (6.2)\\ R \;&=\; D_j\left(\hat{R}-\frac{2}{p^2}\lambda U\mathcal{M}U\right)U^{-1}, \end{aligned}$$

where we put $\hat{R}_2 = \hat{R}\left(\begin{smallmatrix} 0 \\ 1\end{smallmatrix}\right)$ and $\lambda = (0\ \lambda_2) \in \mathbb{Z}^{(n,2)}$.
By the definition of $V_{\alpha,n-\alpha}(p^2)$ there exists $\{\gamma_i\}_i$ such that $\phi|V_{\alpha,n-\alpha}(p^2) = \sum_i \phi|_{k,\mathcal{M}}\gamma_i$. We can take γ_i as a form

$$\gamma_i \;=\; \gamma_{(j,l,\lambda_2,\mu_2)} = \left(\begin{pmatrix} p^{2\,t}D_j^{\;-1} & B_{(j,l)} \\ 0_n & D_j\end{pmatrix}\times\begin{pmatrix} U & 0_2 \\ 0_2 & p^2U^{-1}\end{pmatrix}, [((0\ \lambda_2),(0\ \mu_2)),0_2]\right),$$

where $B_{(j,l)} = \left(\begin{smallmatrix} B^*_{(j,l)} & b_1 \\ {}^tb_3 & b_2\end{smallmatrix}\right)$, $D_j = \left(\begin{smallmatrix} D_j^* & \mathfrak{d} \\ 0 & d_j\end{smallmatrix}\right)$, $\lambda_2 = \left(\begin{smallmatrix} \lambda^* \\ \lambda_3\end{smallmatrix}\right)$, $\mu_2 = \left(\begin{smallmatrix} \mu^* \\ \mu_3\end{smallmatrix}\right)$ with $\left(\begin{smallmatrix} p^{2\,t}D_j^{*\,-1} & B^*_{(j,l)} \\ 0_{n-1} & D_j^*\end{smallmatrix}\right) \in \mathrm{GSp}^+_{n-1}(\mathbb{Z})$, λ^*, $\mu^* \in \mathbb{Z}^{(n-1,1)}$, and d_j, λ_3, $\mu_3 \in \mathbb{Z}$. We set

$$\gamma_i^* := \gamma^*_{(j,l,\lambda^*,\mu^*)} = \left(\begin{pmatrix} p^{2t}D_j^{*-1} & B^*_{(j,l)} \\ 0_{n-1} & D_j^* \end{pmatrix} \times \begin{pmatrix} U & 0_2 \\ 0_2 & p^2U^{-1} \end{pmatrix}, [((0\ \lambda^*),(0\ \mu^*)),0_2] \right).$$

By the definition of Siegel Φ-operator we have

$$\begin{aligned} \Phi\left(\sum_i \phi|_{k,\mathcal{M}}\gamma_i\right)(\tau^*,z^*) &= \Phi(\phi|V_{\alpha,n-\alpha}(p^2))(\tau^*,z^*) \\ &= \sum_{\hat{N},\hat{R}} A_2(\hat{N},\hat{R})e(\hat{N}^*\tau^* + \hat{R}^{*t}z^*), \end{aligned}$$

where $\tau^* \in \mathfrak{H}_{n-1}$, $z^* \in \mathbb{C}^{(n-1,2)}$, $\hat{N} = \left(\begin{smallmatrix} \hat{N}^* & 0 \\ 0 & 0 \end{smallmatrix}\right) \in \mathrm{Sym}_n^*$, $\hat{N}^* \in \mathrm{Sym}^*_{n-1}$, $\hat{R} = \left(\begin{smallmatrix} \hat{R}^* \\ 0 \end{smallmatrix}\right) \in \mathbb{Z}^{(n,2)}$ and $\hat{R}^* \in \mathbb{Z}^{(n-1,2)}$.
Hence we need to calculate $A_2(\hat{N},\hat{R})$ for $\hat{N} = \left(\begin{smallmatrix} \hat{N}^* & 0 \\ 0 & 0 \end{smallmatrix}\right)$ and $\hat{R} = \left(\begin{smallmatrix} \hat{R}^* \\ 0 \end{smallmatrix}\right) \in \mathbb{Z}^{(n,2)}$.
From the identity (6.1) we need to calculate

$$\sum_j \det(D_j)^{-k} \sum_{\lambda_2 \in (\mathbb{Z}/p\mathbb{Z})^{(n,1)}} A_1(N,R) \sum_l e(NB_{(j,l)}D_j^{-1}). \tag{6.3}$$

We remark that the value $A_1(N,R)$ depends on the choice of $\hat{N}$, $\hat{R}$, D_j and λ_3. Under the conditions $\hat{N} \in \mathrm{Sym}_n^*$ and $\hat{R} \in \mathbb{Z}^{(n,2)}$ and by the identity (6.2) we can assume $d\lambda_3 \in p\mathbb{Z}$, since $A_1(N,R) = 0$ unless $N \in \mathrm{Sym}_n^*$. It is known that the value $A_1(N,R)$ depends only on $4N - R\mathcal{M}^{-1t}R$ and on $R \mod 2\mathcal{M}$. We now have

$$4N - R\mathcal{M}^{-1t}R \quad = \quad \frac{1}{p^2}D_j\left(4\hat{N} - p^2\hat{R}U^{-1}\mathcal{M}^{-1}U^{-1t}\hat{R}\right){}^tD_j.$$

We set

$$R' \quad = \quad D_j\left(\hat{R} - \frac{2}{p}(0\ \lambda_2)\mathcal{M}U\right)U^{-1} + \frac{2}{p}\begin{pmatrix} 0 & 0 \\ 0 & d_j\lambda_3 \end{pmatrix}\mathcal{M}$$

and

$$N' \quad = \quad \frac{1}{4p^2}D_j\left(4\hat{N} - p^2\hat{R}U^{-1}\mathcal{M}^{-1}U^{-1t}\hat{R}\right){}^tD_j + \frac{1}{4}R'\mathcal{M}^{-1t}R'.$$

We remark that the last row of R' is zero, and the last row and the last column of N' are also zero. Because $4N - R\mathcal{M}^{-1t}R = 4N' - R'\mathcal{M}^{-1t}R'$ and because $R - R' \in 2\mathbb{Z}^{(n-1,2)}\mathcal{M}$, we have $A_1(N,R) = A_1(N',R')$. We write $N' = \left(\begin{smallmatrix} N'^* & \\ & 0 \end{smallmatrix}\right)$ with $N'^* \in \mathrm{Sym}^*_{n-1}$.
We have

$$R' \quad = \quad D_j\left(\hat{R} - \frac{2}{p}(0\ \lambda_2')\mathcal{M}U\right)U^{-1},$$

where $\lambda_2' = \left(\begin{smallmatrix} \lambda^* - D_j^{*-1}\mathfrak{d}\lambda_3 \\ 0 \end{smallmatrix}\right) \in \mathbb{Q}^{(n,1)}$. Now we will show $\lambda^* - D_j^{*-1}\mathfrak{d}\lambda_3 \in \mathbb{Z}^{(n-1,1)}$ if $\sum_l e(NB_{(j,l)}D_j^{-1}) \neq 0$ in the sum (6.3).
We remark $d_j = 1$, p or p^2. Because $p^2 D_j^{-1} \in \mathbb{Z}^{(n,n)}$ we have $p^2 D_j^{*-1}\mathfrak{d} d_j^{-1} \in \mathbb{Z}^{(n-1,1)}$. If $d_j = 1$, then we can take $\mathfrak{d} = 0 \in \mathbb{Z}^{(n-1,1)}$ as a representative. If $d_j = p^2$, then $D_j^{*-1}\mathfrak{d} \in \mathbb{Z}^{(n-1,1)}$. We now assume $d_j = p$. Then $pD_j^{*-1}\mathfrak{d} \in \mathbb{Z}^{(n-1,1)}$. By using the identity ${}^tB_{(j,l)}D_j = {}^tD_jB_{(j,l)}$ we have

$$\begin{aligned} & e(NB_{(j,l)}D_j^{-1}) \\ = \;& e(N'B_{(j,l)}D_j^{-1} - \frac{d_j\lambda_3}{2p} R \left(\begin{smallmatrix} 0 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & 1 \end{smallmatrix}\right) B_{(j,l)}D_j^{-1} - \frac{d_j\lambda_3}{2p} \begin{pmatrix} 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ 0 & 1 \end{pmatrix} {}^tRB_{(j,l)}D_j^{-1} \\ & - \frac{d_j^2\lambda_3^2}{p^2} \left(\begin{smallmatrix} 0_{n-1} & \\ & 1 \end{smallmatrix}\right) B_{(j,l)}D_j^{-1}) \\ = \;& e(N'^*B^*_{(j,l)}D_j^{*-1})\, e\left(-\frac{d_j\lambda_3}{p^2} (\hat{R}_2^* - 2\lambda^* - D_j^{*-1}\mathfrak{d}_{\mathfrak{j}}\lambda_3)^t b_3 \right) e\left(\frac{d_j\lambda_3^2}{p^2} b_2 \right). \end{aligned}$$

Hence, if $d_j = p$, then $\displaystyle\sum_{b_2 \mod p} e\left(\frac{d_j\lambda_3^2}{p^2}b_2\right)$ is zero unless $\lambda_3 \equiv 0 \mod p$. Thus, for any $d_j \in \{1, p, p^2\}$, we conclude $\sum_l e(NB_{(j,l)}D_j^{-1}) = 0$ in the sum (6.3) unless $D_j^{*-1}\mathfrak{d}\lambda_3 \in \mathbb{Z}^{(n-1,1)}$. Hence $\lambda^* - D_j^{*-1}\mathfrak{d}\lambda_3 \in \mathbb{Z}^{(n-1,1)}$ and $\lambda_2' \in \mathbb{Z}^{(n,1)}$, if $\sum_l e(NB_{(j,l)}D_j^{-1}) \neq 0$.
Therefore there exists a set of complex numbers $\{C_{\gamma_i}\}_i := \{C_{\gamma_i,k,\mathcal{M}}\}_i$ which satisfies

$$\Phi\left(\sum_i \phi|_{k,\mathcal{M}}\gamma_i\right) = \sum_i C_{\gamma_i^*}\Phi(\phi)|_{k,\mathcal{M}}\gamma_i^*.$$

By a well-known argument we have $\sum_i C_{\gamma_i^*}\gamma_i^*\gamma = \sum_i C_{\gamma_i^*}\gamma_i^*$ for any $\gamma \in \Gamma^J_{n-1,2}$. Hence there exists an index-shift map $V_{\alpha,n-\alpha}(p^2)^*$ which satisfies the identity $\Phi(\phi|V_{\alpha,n-\alpha}(p^2)) = \Phi(\phi)|V_{\alpha,n-\alpha}(p^2)^*$.
For a fixed α $(0 \leq \alpha \leq n)$ the index-shift map $V_{\alpha,n-\alpha}(p^2)^*$ is a linear combination of $V_{\beta,n-1-\beta}(p^2)$ $(\beta = 0, ..., n-1)$. We need to determine these coefficients of the linear combination. This calculation is similar to the case of Siegel modular forms [Kr 86, page 325]. We leave the details to the reader.

$\square$

Now for integers l $(2 \leq l)$, β $(0 \leq \beta \leq l-1)$ and α $(0 \leq \alpha \leq l)$, we put

$$b_{\beta,\alpha} \;:=\; b_{\beta,\alpha,l,p}(X) = \begin{cases} (p^{l+1-\alpha} - p^{-l-1+\alpha})p^{\frac{1}{2}} & \text{if } \beta = \alpha - 2 \\ (X + X^{-1})p & \text{if } \beta = \alpha - 1 \\ p^{-l+\alpha+\frac{3}{2}} & \text{if } \beta = \alpha \\ 0 & \text{otherwise} \end{cases}$$

and set a matrix

$$B_{l,l+1}(X) \quad := \quad (b_{\beta,\alpha})_{\substack{\beta=0,\dots,l-1\\ \alpha=0,\dots,l}} = \begin{pmatrix} b_{0,0} & \cdots & b_{0,l} \\ \vdots & \ddots & \vdots \\ b_{l-1,0} & \cdots & b_{l-1,l} \end{pmatrix}$$

with entries in $\mathbb{C}[X+X^{-1}]$. For any $\phi \in J_{k,\mathcal{M}}^{(l)}$, due to Proposition 6.1, we obtain

$$\begin{aligned} &\Phi(\phi)|(V_{0,l}(p^2)^*, \cdots, V_{l,0}(p^2)^*) \\ &= p^{-k+l+\frac{1}{2}} \left(\Phi(\phi)|(V_{0,l-1}(p^2), \cdots, V_{l-1,0}(p^2))\right) B_{l,l+1}(p^{k-l-\frac{1}{2}}). \end{aligned} \tag{6.4}$$

Here $\Phi(\phi)|(V_{0,l}(p^2)^*, \cdots, V_{l,0}(p^2)^*)$ denotes the row vector

$$\Phi(\phi)|(V_{0,l}(p^2)^*, \cdots, V_{l,0}(p^2)^*) \quad := \quad \left(\Phi(\phi)|V_{0,l}(p^2)^*, ..., \Phi(\phi)|V_{l,0}(p^2)^*\right).$$

Let $J_{k-\frac{1}{2},m}^{(n)*}$ be the subspace of $J_{k-\frac{1}{2},m}^{(n)}$ introduced in §4.4.

Corollary 6.2. *For any Jacobi form $\phi \in J_{k-\frac{1}{2},m}^{(n)*}$ and for any prime p, we have*

$$\Phi(\phi|\tilde{V}_{\alpha,n-\alpha}(p^2)) \quad = \quad \Phi(\phi)|\tilde{V}_{\alpha,n-\alpha}(p^2)^*,$$

where $\tilde{V}_{\alpha,n-\alpha}(p^2)^*$ is a map $\tilde{V}_{\alpha,n-\alpha}(p^2)^* : J_{k-\frac{1}{2},m}^{(n-1)*} \to J_{k-\frac{1}{2},mp^2}^{(n-1)}$ given by

$$\begin{aligned} \widetilde{V}_{\alpha,n-\alpha}(p^2)^* \quad = \quad & p^{k-n-\frac{1}{2}} \Big\{ p^{-n+\alpha} \widetilde{V}_{\alpha,n-\alpha-1}(p^2) \\ & + (p^{-k+n+\frac{1}{2}} + p^{k-n-\frac{1}{2}}) \widetilde{V}_{\alpha-1,n-\alpha}(p^2) \\ & + (p^{n+1-\alpha} - p^{-n-1+\alpha}) \widetilde{V}_{\alpha-2,n-\alpha+1}(p^2) \Big\}. \end{aligned}$$

Proof. By a straightforward calculation we get the fact that $\iota_{\mathcal{M}}$ and Φ is commutative. The rest of the proof of this corollary follows from Proposition 6.1 and Proposition 4.4. $\square$

Let $\tilde{\mathcal{H}}_{p^2}^{(m)}$ be the local Hecke ring and let R_m be the subring of a polynomial ring both defined in §2.9. The isomorphism $\Psi_m : \tilde{\mathcal{H}}_{p^2}^{(m)} \cong R_m$ has been obtained in [Zh 83, Zh 84] (see §2.9).

Proposition 6.3. *Let p be an odd prime. For any $m \geq 2$, the image of generators $K_\alpha^{(m)}$ of $\tilde{\mathcal{H}}_{p^2}^{(m)}$ by Ψ_m are expressed as a vector*

$$\begin{aligned}
&\left(\Psi_m(K_0^{(m)}), \Psi_m(K_1^{(m)}), \cdots, \Psi_m(K_m^{(m)})\right) \\
&= p^{-\frac{3}{2}(m-1)}\ z_0^2 z_1 \cdots z_m (p^{-1},\ z_1 + z_1^{-1}) \begin{pmatrix} 1 & 0 \\ 0 & p^{\frac{1}{2}} \end{pmatrix}^{-1} \\
&\qquad \times \left\{ \prod_{l=2}^{m} B_{l,l+1}(z_l) \right\}\ diag(1, p^{\frac{1}{2}}, ..., p^{\frac{m}{2}}).
\end{aligned} \tag{6.5}$$

Here $B_{l,l+1}(X)$ is the $l \times (l+1)$-matrix introduced in above, and where

$$\prod_{l=2}^{m} B_{l,l+1}(z_l) = B_{2,3}(z_2) B_{3,4}(z_3) \cdots B_{m,m+1}(z_m)$$

is a $2 \times (m+1)$ matrix with entries in $\mathbb{C}[z_2^{\pm}, \cdots, z_m^{\pm}]$. We remark that

$$\Psi_m(K_0^{(m)}) = p^{-\frac{m(m+1)}{2}}\ z_0^2 z_1 \cdots z_m.$$

Proof. Let k be an even integer and let $F \in M_{k-\frac{1}{2}}(\Gamma_0^{(m)}(4))$ be a Siegel modular form such that $\Phi^S(F) \not\equiv 0$. Here Φ^S denotes the Siegel Φ-operator for Siegel modular forms. Let $T \in \tilde{H}_{p^2}^{(m)}$ and let $f_T(z_0, ..., z_m) := \Psi_m(T) \in R_m$. Then $f_T(z_0, ... z_{m-1}, p^{k-m-\frac{1}{2}}) \in R_{m-1}$ and $\Psi_{m-1}^{-1}(f_T(z_0, ..., z_{m-1}, p^{k-m-\frac{1}{2}})) \in \tilde{H}_{p^2}^{(m-1)}$. It is known by Oh-Koo-Kim [OKK 89, Theorem 5.1] that

$$\Phi^S(F|T) \quad = \quad \Phi^S(F)|\Psi_{m-1}^{-1}(f_T(z_0, ..., z_{m-1}, p^{k-m-\frac{1}{2}})). \tag{6.6}$$

Let $\phi \in J_{k-\frac{1}{2},a}^{(m)}$ be a Jacobi form with index $a \in \mathbb{Z}$ such that $\Phi(\phi) \not\equiv 0$. Here Φ is the Siegel Φ-operator. If k is large enough, then there exists such ϕ. Due to Corollary 6.2 we have

$$\Phi(\phi|\tilde{V}_{\alpha,m-\alpha}(p^2)) \quad = \quad \Phi(\phi)|\tilde{V}_{\alpha,m-\alpha}(p^2)^*.$$

Let $\mathbb{W} : J_{k-\frac{1}{2},a}^{(m)} \to M_{k-\frac{1}{2}}^{+}(\Gamma_0^{(m)}(4))$ be the Witt operator which is defined by

$$\mathbb{W}(\phi)(\tau) := \phi(\tau, 0)$$

for any $\phi(\tau, z) \in J_{k-\frac{1}{2},a}^{(m)}$. By a straightforward calculation, for any $\phi \in J_{k-\frac{1}{2},a}^{(m)}$ we have

$$\mathbb{W}(\phi|\tilde{V}_{\alpha,m-\alpha}(p^2)) \quad = \quad \mathbb{W}(\phi)|\tilde{T}_{\alpha,m-\alpha}(p^2)$$

and

$$\mathbb{W}(\Phi(\phi)) \quad = \quad \Phi^S(\mathbb{W}(\phi)).$$

We set

$$\begin{aligned}\widetilde{T}_{\alpha,m-\alpha}(p^2)^* &= p^{k-m-\frac{1}{2}}\Big\{p^{-m+\alpha}\widetilde{T}_{\alpha,m-\alpha-1}(p^2)\\ &+(p^{-k+m+\frac{1}{2}}+p^{k-m-\frac{1}{2}})\widetilde{T}_{\alpha-1,m-\alpha}(p^2)\\ &+(p^{m+1-\alpha}-p^{-m-1+\alpha})\widetilde{T}_{\alpha-2,m-\alpha+1}(p^2)\Big\}.\end{aligned}$$

If we put $F=\mathbb{W}(\phi)$, then

$$\begin{aligned}\Phi^S(F|\tilde{T}_{\alpha,m-\alpha}(p^2)) &= \Phi^S(\mathbb{W}(\phi|\tilde{V}_{\alpha,m-\alpha}(p^2)))\\ &= \mathbb{W}(\Phi(\phi|\tilde{V}_{\alpha,m-\alpha}(p^2)))\\ &= \mathbb{W}(\Phi(\phi)|\tilde{V}_{\alpha,m-\alpha}(p^2)^*)\\ &= \Phi^S(F)|\tilde{T}_{\alpha,m-\alpha}(p^2)^*.\end{aligned}$$

Hence if we put $T=\tilde{T}_{\alpha,m-\alpha}(p^2)$ in (6.6) we have

$$\begin{aligned}f_T(z_0,...,z_{m-1},p^{k-m-\frac{1}{2}}) &= p^{k-m-\frac{1}{2}}\Big\{p^{-m+\alpha}\Psi_{m-1}(K_\alpha^{(m-1)})\\ &+(p^{-k+m+\frac{1}{2}}+p^{k-m-\frac{1}{2}})\Psi_{m-1}(K_{\alpha-1}^{(m-1)})\\ &+(p^{m+1-\alpha}-p^{-m-1+\alpha})\Psi_{m-1}(K_{\alpha-2}^{(m-1)})\Big\}.\end{aligned}$$

Since this identity is true for infinitely many k, we have

$$\begin{aligned}\Psi_m(K_\alpha^{(m)}) &= f_T(z_0,...,z_{m-1},z_m)\\ &= z_m\Big\{p^{-m+\alpha}\Psi_{m-1}(K_\alpha^{(m-1)})\\ &+(z_m+z_m^{-1})\Psi_{m-1}(K_{\alpha-1}^{(m-1)})\\ &+(p^{m+1-\alpha}-p^{-m-1+\alpha})\Psi_{m-1}(K_{\alpha-2}^{(m-1)})\Big\}.\end{aligned}$$

Hence

$$\begin{aligned}&\Big(\Psi_m(K_0^{(m)}),\Psi_m(K_1^{(m)}),\cdots,\Psi_m(K_m^{(m)})\Big)\\ &=p^{-3/2}\ z_m\ \Big(\Psi_{m-1}(K_0^{(m-1)}),\Psi_{m-1}(K_1^{(m-1)}),\cdots,\Psi_{m-1}(K_{m-1}^{(m-1)})\Big)\\ &\quad\times\operatorname{diag}(1,p^{\frac{1}{2}},...,p^{\frac{m-1}{2}})^{-1}B_{m,m+1}(z_m)\ \operatorname{diag}(1,p^{\frac{1}{2}},...,p^{\frac{m}{2}}).\end{aligned}$$

Moreover, we have

$$\begin{aligned}\Psi_1\Big(K_0^{(1)}\Big) &= p^{-1}z_0^2z_1,\\ \Psi_1\Big(K_1^{(1)}\Big) &= z_0^2(1+z_1^2)\end{aligned}$$

by the definition of Ψ_1. This proposition follows from these identities and the recursion with respect to m. $\square$

7 Maass relation for generalized Cohen-Eisenstein series

We put a $2\times(n+1)$-matrix

$$\begin{aligned} A^p_{2,n+1}(X) &:= \prod_{l=2}^{n} B_{l,l+1}(p^{\frac{n+2}{2}-l}X) \\ &= B_{2,3}(p^{\frac{n+2}{2}-2}X)B_{3,4}(p^{\frac{n+2}{2}-3}X)\cdots B_{n,n+1}(p^{\frac{n+2}{2}-n}X), \end{aligned}$$

where $B_{l,l+1}(X)$ is the $l\times(l+1)$-matrix introduced in §6.

Lemma 7.1. *All components of the matrix $A^p_{2,2n-1}(X)$ belong to $\mathbb{C}[X+X^{-1}]$.*

Proof. We assume p is an odd prime. Let R_m be the symbol introduced in §2.9. Because $\Psi_{2n-2}(K^{(2n-2)}_\alpha)$ belongs to R_{2n-2} and because of Proposition 6.3 we have relations $B_{l,l+1}(z_l) = B_{l,l+1}(z_l^{-1})$ $(l=2,...,2n-2)$ and

$$\begin{aligned} &B_{2,3}(z_2)B_{3,4}(z_3)\cdots B_{2n-2,2n-1}(z_{2n-2}) \\ &\quad = B_{2,3}(z_{2n-2})B_{3,4}(z_{2n-3})\cdots B_{2n-2,2n-1}(z_2). \end{aligned}$$

Hence

$$\begin{aligned} A^p_{2,2n-1}(X) &= B_{2,3}(p^{n-2}X)B_{3,4}(p^{n-3}X)\cdots B_{2n-2,2n-1}(p^{-n+2}X) \\ &= B_{2,3}(p^{-n+2}X^{-1})B_{3,4}(p^{-n+3}X^{-1})\cdots B_{2n-2,2n-1}(p^{n-2}X^{-1}) \\ &= B_{2,3}(p^{n-2}X^{-1})B_{3,4}(p^{n-3}X^{-1})\cdots B_{2n-2,2n-1}(p^{-n+2}X^{-1}) \\ &= A^p_{2,2n-1}(X^{-1}). \end{aligned}$$

The relation $A^p_{2,2n-1}(X) = A^p_{2,2n-1}(X^{-1})$ holds for infinitely many p. Hence if we regard that the components of the matrix $A^p_{2,2n-1}(X)$ are Laurent-polynomials of variables X and $p^{1/2}$, then we obtain $A^p_{2,2n-1}(X) = A^p_{2,2n-1}(X^{-1})$. Hence we have also $A^p_{2,2n-1}(X) = A^p_{2,2n-1}(X^{-1})$ for $p=2$. $\square$

Let $\mathcal{M}$, m, D_0 and f be the symbols used in the previous sections, it means that $\mathcal{M} = \begin{pmatrix} * & * \\ * & 1 \end{pmatrix}$ is a 2×2 half-integral symmetric-matrix, $m=\det(2\mathcal{M})$, D_0 is the discriminant of $\mathbb{Q}(\sqrt{-m})$ and f is a non-negative integer which satisfies $m = D_0 f^2$.

For any prime p we set

$$\begin{pmatrix} a_{0,m,p,k} \\ a_{1,m,p,k} \\ a_{2,m,p,k} \end{pmatrix} := \begin{cases} \begin{pmatrix} p^{-3k+4} \\ 0 \\ p^{-k+1} \end{pmatrix} & \text{if } p|f, \\ \begin{pmatrix} 0 \\ p^{-3k+4} + p^{-2k+2}\left(\frac{-m}{p}\right) \\ p^{-k+1} - p^{-2k+2}\left(\frac{-m}{p}\right) \end{pmatrix} & \text{if } p\not| f. \end{cases}$$

LEMMA 7.2. *For the Jacobi-Eisenstein series $E^{(1)}_{k,\mathcal{M}}$ of weight k of index $\mathcal{M}$ of degree 1, we have the identity*

$$\begin{aligned} & E^{(1)}_{k,\mathcal{M}}|(V_{0,1}(p^2), V_{1,0}(p^2)) \\ = & \left(E^{(1)}_{k,\mathcal{M}\left[X^{-1}\left(\begin{smallmatrix} p & 0 \\ 0 & 1\end{smallmatrix}\right)^{-1}\right]} |U_{\left(\begin{smallmatrix} p & 0 \\ 0 & 1\end{smallmatrix}\right) X \left(\begin{smallmatrix} p & 0 \\ 0 & 1\end{smallmatrix}\right)}, E^{(1)}_{k,\mathcal{M}}|U_{\left(\begin{smallmatrix} p & 0 \\ 0 & 1\end{smallmatrix}\right)}, E^{(1)}_{k,\mathcal{M}[\left(\begin{smallmatrix} p & 0 \\ 0 & 1\end{smallmatrix}\right)]} \right) \\ & \times \begin{pmatrix} 0 & a_{0,m,p,k} \\ p^{-2k+2} & a_{1,m,p,k} \\ 0 & a_{2,m,p,k} \end{pmatrix}, \end{aligned}$$

where $X = \left(\begin{smallmatrix} 1 & 0 \\ x & 1\end{smallmatrix}\right) \in \mathbb{Z}^{(2,2)}$ is a matrix such that $\mathcal{M}[X^{-1}\left(\begin{smallmatrix} p & 0 \\ 0 & 1\end{smallmatrix}\right)^{-1}] \in Sym_2^+$. Here, if $p\not| f$, there does not exist such matrix X and we regard $E^{(1)}_{k,\mathcal{M}\left[X^{-1}\left(\begin{smallmatrix} p & 0 \\ 0 & 1\end{smallmatrix}\right)^{-1}\right]}$ as zero.

Proof. From Proposition 5.3 and due to (5.2), (5.3) in the proof of Proposition 5.5, we have

$$\begin{aligned} & E^{(1)}_{k,\mathcal{M}}|V_{0,1}(p^2) = \tilde{K}^0_{0,1} \\ = & \; p^{-2k+1} \sum_{M \in \Gamma(\delta_{0,1})\backslash\Gamma_1} \sum_{\lambda_2 \in L^*_{0,1}} \{1|_{k,\mathcal{M}}([(\lambda_2,0),0],M)\}\left(\tau, z\left(\begin{smallmatrix} p & 0 \\ 0 & 1\end{smallmatrix}\right)\right) \\ & \times \sum_{u_2 \in \mathbb{Z}/p\mathbb{Z}} G^{1,1}_{\mathcal{M}}(\lambda_2 + (0,u_2)) \\ = & \; p^{-2k+2} \sum_{M \in \Gamma^{(1)}_\infty\backslash\Gamma_1} \sum_{\lambda_2 \in \mathbb{Z}^{(1,2)}} \{1|_{k,\mathcal{M}}([(\lambda_2,0),0],M)\}\left(\tau, z\left(\begin{smallmatrix} p & 0 \\ 0 & 1\end{smallmatrix}\right)\right) \\ = & \; p^{-2k+2} E^{(1)}_{k,\mathcal{M}}|U_{\left(\begin{smallmatrix} p & 0 \\ 0 & 1\end{smallmatrix}\right)}. \end{aligned}$$

From Proposition 5.3 we also have

$$E^{(1)}_{k,\mathcal{M}}|V_{1,0}(p^2) = \tilde{K}^0_{1,1} + \tilde{K}^1_{0,1} + \tilde{K}^0_{0,0}.$$

Here

$$\begin{aligned}
\tilde{K}^0_{1,1} &= p^{-k+1} \sum_{M\in\Gamma(\delta_{1,1})\backslash\Gamma_1} \sum_{\lambda_1\in L^*_{1,1}} \{1|_{k,\mathcal{M}}([(\lambda_1,0),0],M)\}(\tau, z\left(\begin{smallmatrix} p & 0\\ 0 & 1\end{smallmatrix}\right)) \\
&= p^{-k+1} \sum_{M\in\Gamma^{(1)}_\infty\backslash\Gamma_1} \sum_{\lambda_1\in p\mathbb{Z}\times\mathbb{Z}} \{1|_{k,\mathcal{M}}([(\lambda_1,0),0],M)\}(\tau, z\left(\begin{smallmatrix} p & 0\\ 0 & 1\end{smallmatrix}\right)) \\
&= p^{-k+1} \sum_{M\in\Gamma^{(1)}_\infty\backslash\Gamma_1} \sum_{\lambda\in\mathbb{Z}^{(1,2)}} \{1|_{k,\mathcal{M}}([(\lambda\left(\begin{smallmatrix} p & \\ & 1\end{smallmatrix}\right),0),0],M)\}(\tau, z\left(\begin{smallmatrix} p & 0\\ 0 & 1\end{smallmatrix}\right)) \\
&= p^{-k+1} \sum_{M\in\Gamma^{(1)}_\infty\backslash\Gamma_1} \sum_{\lambda\in\mathbb{Z}^{(1,2)}} \left\{1|_{k,\mathcal{M}[\left(\begin{smallmatrix} p & \\ & 1\end{smallmatrix}\right)]}([(\lambda,0),0],M)\right\}(\tau, z) \\
&= p^{-k+1} E^{(1)}_{k,\mathcal{M}[\left(\begin{smallmatrix} p & \\ & 1\end{smallmatrix}\right)]}(\tau,z).
\end{aligned}$$

Now we shall calculate $\tilde{K}^1_{0,1}$. First, due to Lemma 5.4 we have

$$\sum_{u_2\in\mathbb{Z}/p\mathbb{Z}} G^{1,0}_{\mathcal{M}}(\lambda+(0,u_2)) = \begin{cases} 0 & \text{if } \lambda\in\mathbb{Z}^{(1,2)}\left(\begin{smallmatrix} p & 0\\ 0 & 1\end{smallmatrix}\right), \\ \left(\frac{-m}{p}\right)p & \text{if } \lambda\notin\mathbb{Z}^{(1,2)}\left(\begin{smallmatrix} p & 0\\ 0 & 1\end{smallmatrix}\right)\end{cases}$$

for any $\lambda\in\mathbb{Z}^{(1,2)}$. Thus

$$\begin{aligned}
&\tilde{K}^1_{0,1} \\
&= p^{-2k+2} \sum_{M\in\Gamma(\delta_{0,1})\backslash\Gamma_1} \sum_{\lambda_2\in L^*_{0,1}} \{1|_{k,\mathcal{M}}([(\lambda_2,0),0],M)\}(\tau, z\left(\begin{smallmatrix} p & 0\\ 0 & 1\end{smallmatrix}\right)) \\
&\qquad \times \sum_{u_2\in\mathbb{Z}/p\mathbb{Z}} G^{1,0}_{\mathcal{M}}(\lambda_2+(0,u_2)) \\
&= -p^{-2k+2}\left(\frac{-m}{p}\right) \sum_{M\in\Gamma^{(1)}_\infty\backslash\Gamma_1} \sum_{\lambda\in\mathbb{Z}^{(1,2)}} \{1|_{k,\mathcal{M}}([(\lambda\left(\begin{smallmatrix} p & 0\\ 0 & 1\end{smallmatrix}\right),0),0],M)\}(\tau, z\left(\begin{smallmatrix} p & 0\\ 0 & 1\end{smallmatrix}\right)) \\
&\qquad + p^{-2k+2}\left(\frac{-m}{p}\right) \sum_{M\in\Gamma^{(1)}_\infty\backslash\Gamma_1} \sum_{\lambda\in\mathbb{Z}^{(1,2)}} \{1|_{k,\mathcal{M}}([(\lambda,0),0],M)\}(\tau, z\left(\begin{smallmatrix} p & 0\\ 0 & 1\end{smallmatrix}\right)) \\
&= -p^{-2k+2}\left(\frac{-m}{p}\right) E^{(1)}_{k,\mathcal{M}\left[\left(\begin{smallmatrix} p & 0\\ 0 & 1\end{smallmatrix}\right)\right]}(\tau,z) + p^{-2k+2}\left(\frac{-m}{p}\right) E_{k,\mathcal{M}}(\tau, z\left(\begin{smallmatrix} p & 0\\ 0 & 1\end{smallmatrix}\right)) \\
&= -p^{-2k+2}\left(\frac{-m}{p}\right) E^{(1)}_{k,\mathcal{M}\left[\left(\begin{smallmatrix} p & 0\\ 0 & 1\end{smallmatrix}\right)\right]} + p^{-2k+2}\left(\frac{-m}{p}\right) E_{k,\mathcal{M}}|U_{\left(\begin{smallmatrix} p & 0\\ 0 & 1\end{smallmatrix}\right)}.
\end{aligned}$$

We shall calculate $\tilde{K}^0_{0,0}$. Due to (5.2) and due to (5.3) we have

$$L^*_{0,0} = \begin{cases} \mathbb{Z}^{(1,2)}\left(\begin{smallmatrix} p & 0\\ 0 & 1\end{smallmatrix}\right)^{-1}\,{}^tX^{-1} & \text{if } p|f, \\ \mathbb{Z}^{(1,2)} & \text{if } p\nmid f.\end{cases}$$

Thus if $p|f$, then

$$
\begin{aligned}
\tilde{K}^0_{0,0} &= p^{-3k+4} \sum_{M \in \Gamma(\delta_{0,0}) \backslash \Gamma_1} \sum_{\lambda_3 \in \mathbb{Z}^{(1,2)} \left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)^{-1} {}^t X^{-1}} \left\{ 1|_{k,\mathcal{M}}([(\lambda_3,0),0],M) \right\} \left(\tau, z \left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)\right) \\
&= p^{-3k+4} \sum_{M \in \Gamma^{(1)}_\infty \backslash \Gamma_1} \sum_{\lambda_3 \in \mathbb{Z}^{(1,2)} \left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)^{-1} {}^t X^{-1}} \\
&\quad \times \left\{ 1|_{k,\mathcal{M}[X^{-1}\left(\begin{smallmatrix} p & \\ & 1 \end{smallmatrix}\right)^{-1}]}([(\lambda_3 {}^t X \left(\begin{smallmatrix} p & \\ & 1 \end{smallmatrix}\right),0),0],M) \right\} \left(\tau, z \left(\begin{smallmatrix} p & \\ & 1 \end{smallmatrix}\right) {}^t X \left(\begin{smallmatrix} p & \\ & 1 \end{smallmatrix}\right)\right) \\
&= p^{-3k+4} E^{(1)}_{k,\mathcal{M}[X^{-1}\left(\begin{smallmatrix} p & \\ & 1 \end{smallmatrix}\right)^{-1}]}\left(\tau, z \left(\begin{smallmatrix} p & \\ & 1 \end{smallmatrix}\right) {}^t X \left(\begin{smallmatrix} p & \\ & 1 \end{smallmatrix}\right)\right),
\end{aligned}
$$

and if $p \nmid f$, then

$$
\begin{aligned}
\tilde{K}^0_{0,0} &= p^{-3k+4} \sum_{M \in \Gamma(\delta_{0,0}) \backslash \Gamma_1} \sum_{\lambda_3 \in \mathbb{Z}^{(1,2)}} \left\{ 1|_{k,\mathcal{M}}([(\lambda_3,0),0],M) \right\} \left(\tau, z \left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)\right) \\
&= p^{-3k+4} E^{(1)}_{k,\mathcal{M}}\left(\tau, z \left(\begin{smallmatrix} p & \\ & 1 \end{smallmatrix}\right)\right).
\end{aligned}
$$

Hence we obtain the formula for $\tilde{K}^0_{0,0}$.
Because $E^{(1)}_{k,\mathcal{M}}|V_{1,0}(p^2) = \tilde{K}^0_{1,1} + \tilde{K}^1_{0,1} + \tilde{K}^0_{0,0}$, we conclude the lemma. $\square$

LEMMA 7.3. *The three forms* $E^{(1)}_{k,\mathcal{M}\left[X^{-1}\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)^{-1}\right]}|U_{\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right) X \left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)}$, $E^{(1)}_{k,\mathcal{M}}|U_{\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)}$ *and* $E^{(1)}_{k,\mathcal{M}\left[\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)\right]}$ *are linearly independent.*

Proof. We first assume that $\mathcal{M} \in \mathrm{Sym}^+_g$ is a positive-definite half-integral symmetric matrix of size g. Let

$$
E^{(1)}_{k,\mathcal{M}}(\tau,z) = \sum_{\substack{n \in \mathbb{Z}, R \in \mathbb{Z}^{(1,g)} \\ 4n - R\mathcal{M}^{-1} {}^t R \geq 0}} C_{k,\mathcal{M}}(n,R)\, e(n\tau + R^t z)
$$

be the Fourier expansion of the Jacobi-Eisenstein series $E^{(1)}_{k,\mathcal{M}}$. For any pair $n \in \mathbb{Z}$ and $R \in \mathbb{Z}^{(1,g)}$ which satisfy $4n - R\mathcal{M}^{-1t}R > 0$, we now show that $C_{k,\mathcal{M}}(n,R) \neq 0$. The Fourier coefficients of Jacobi-Eisenstein series of degree 1 of integer index have been calculated in [E-Z 85, pp.17–22]. If $4n - R\mathcal{M}^{-1t}R > 0$, by an argument similar to [E-Z 85] we have

$$
C_{k,\mathcal{M}}(n,R) = \frac{(-1)^{\frac{k}{2}} \pi^{k-\frac{g}{2}}}{2^{k-2}\Gamma(k-\frac{g}{2})} \frac{(4n - R\mathcal{M}^{-1t}R)^{k-\frac{g}{2}-1}}{\det(\mathcal{M})^{\frac{1}{2}}} \zeta(k-g)^{-1} \sum_{a=1}^{\infty} \frac{N_a(Q)}{a^{k-1}},
$$

where $N_a(Q) := \left| \left\{ \lambda \in (\mathbb{Z}/a\mathbb{Z})^{(1,g)} \mid \lambda\mathcal{M}^t\lambda + R^t\lambda + n \equiv 0 \mod a \right\} \right|$. Hence we conclude $C_{k,\mathcal{M}}(n,R) \neq 0$.

We now assume $\mathcal{M} = \begin{pmatrix} l & \frac{r}{2} \\ \frac{r}{2} & 1 \end{pmatrix} \in \mathrm{Sym}_2^+$. The (n,R)-th Fourier coefficient of two Jacobi forms $E^{(1)}_{k,\mathcal{M}}|U_{\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)}$ and $E^{(1)}_{k,\mathcal{M}\left[\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)\right]}$ are $C_{k,\mathcal{M}}(n, R\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)^{-1})$ and $C_{k,\mathcal{M}\left[\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)\right]}(n,R)$, respectively. If $R \not\in \mathbb{Z}^{(1,2)}\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)$, then $C_{k,\mathcal{M}}(n, R\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)^{-1}) = 0$ and $C_{k,\mathcal{M}\left[\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)\right]}(n,R) \neq 0$. Hence $E^{(1)}_{k,\mathcal{M}}|U_{\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)}$ and $E^{(1)}_{k,\mathcal{M}\left[\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)\right]}$ are linearly independent. The proof for the linear independence of the three forms of the lemma is similar. We omitted the detail here. □

Proposition 7.4. *We obtain the identity*

$$\begin{aligned}
&E^{(n)}_{k,\mathcal{M}}|\left(V_{0,n}(p^2), \cdots, V_{n,0}(p^2)\right) \\
&= \left(E^{(n)}_{k,\mathcal{M}\left[X^{-1}\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)^{-1}\right]}|U_{\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)X\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)}, E^{(n)}_{k,\mathcal{M}}|U_{\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)}, E^{(n)}_{k,\mathcal{M}\left[\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)\right]}\right) \\
&\quad \times \begin{pmatrix} 0 & a_{0,m,p,k} \\ p^{-2k+2} & a_{1,m,p,k} \\ 0 & a_{2,m,p,k} \end{pmatrix} p^{(-k+\frac{1}{2}n+\frac{3}{2})(n-1)} A^p_{2,n+1}(p^{k-\frac{n+2}{2}-\frac{1}{2}}),
\end{aligned}$$

where the $2 \times (n+1)$*-matrix* $A^p_{2,n+1}(p^{k-\frac{n+2}{2}-\frac{1}{2}})$ *is introduced in the beginning of this section.*

Proof. Let Φ be the Siegel Φ-operator introduced in §6. From the definition of Jacobi-Eisenstein series, we have $\Phi(E^{(l)}_{k,\mathcal{M}}) = E^{(l-1)}_{k,\mathcal{M}}$.
From the identity (6.4) in §6 and from Lemma 7.2, we obtain

$$\begin{aligned}
&\Phi^{n-1}(E^{(n)}_{k,\mathcal{M}}|(V_{0,n}(p^2), ..., V_{n,0}(p^2))) \\
&= \left(E^{(1)}_{k,\mathcal{M}}|(V_{0,1}(p^2), V_{1,0}(p^2))\right)\left(\prod_{l=2}^{n} p^{-k+l+\frac{1}{2}}\right) B_{2,3}(p^{k-\frac{5}{2}}) \cdots B_{n,n+1}(p^{k-n-\frac{1}{2}}) \\
&= p^{(-k+\frac{1}{2})(n-1)+\frac{n(n+1)}{2}-1}\left(E^{(1)}_{k,\mathcal{M}}|(V_{0,1}(p^2), V_{1,0}(p^2))\right) A^p_{2,n+1}(p^{k-\frac{n+2}{2}-\frac{1}{2}}) \\
&= \left(E^{(1)}_{k,\mathcal{M}\left[X^{-1}\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)^{-1}\right]}|U_{\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)X\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)}, E^{(1)}_{k,\mathcal{M}}|U_{\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)}, E^{(1)}_{k,\mathcal{M}\left[\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)\right]}\right) \qquad (7.1) \\
&\quad \times p^{(-k+\frac{1}{2}n+\frac{3}{2})(n-1)} \begin{pmatrix} 0 & a_{0,m,p,k} \\ p^{-2k+2} & a_{1,m,p,k} \\ 0 & a_{2,m,p,k} \end{pmatrix} A^p_{2,n+1}(p^{k-\frac{n+2}{2}-\frac{1}{2}}).
\end{aligned}$$

From Proposition 5.5 there exists a matrix $M \in \mathbb{C}^{(3,n+1)}$ which satisfies

$$\begin{aligned}
&E^{(n)}_{k,\mathcal{M}}|(V_{0,n}(p^2), ..., V_{n,0}(p^2)) \\
&= \left(E^{(n)}_{k,\mathcal{M}\left[X^{-1}\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)^{-1}\right]}|U_{\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)X\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)}, E^{(n)}_{k,\mathcal{M}}|U_{\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)}, E^{(n)}_{k,\mathcal{M}\left[\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)\right]}\right) M. \qquad (7.2)
\end{aligned}$$

Thus

$$
\begin{aligned}
&\Phi^{n-1}(E^{(n)}_{k,\mathcal{M}}|(V_{0,n}(p^2),...,V_{n,0}(p^2)))\\
&= \left(E^{(1)}_{k,\mathcal{M}\left[X^{-1}\left(\begin{smallmatrix} p & 0\\ 0 & 1\end{smallmatrix}\right)^{-1}\right]}|U_{\left(\begin{smallmatrix} p & 0\\ 0 & 1\end{smallmatrix}\right)X\left(\begin{smallmatrix} p & 0\\ 0 & 1\end{smallmatrix}\right)}, E^{(1)}_{k,\mathcal{M}}|U_{\left(\begin{smallmatrix} p & 0\\ 0 & 1\end{smallmatrix}\right)}, E^{(1)}_{k,\mathcal{M}\left[\left(\begin{smallmatrix} p & 0\\ 0 & 1\end{smallmatrix}\right)\right]}\right) M.
\end{aligned}
$$

From Lemma 7.3 the matrix M is uniquely determined. Therefore, by using the identity (7.1), we have

$$
M = p^{(-k+\frac{1}{2}n+\frac{3}{2})(n-1)} \begin{pmatrix} 0 & a_{0,m,p,k}\\ p^{-2k+2} & a_{1,m,p,k}\\ 0 & a_{2,m,p,k}\end{pmatrix} A^p_{2,n+1}(p^{k-\frac{n+2}{2}-\frac{1}{2}}).
$$

Therefore we conclude that this Proposition follows from the identity (7.2). $\square$

We recall that the form $e^{(n)}_{k,\mathcal{M}}$ is the $\mathcal{M}$-th Fourier-Jacobi coefficient of Siegel-Eisenstein series $E^{(n+2)}_k$ of weight k of degree $n+2$.

PROPOSITION 7.5. *We obtain the identity*

$$
\begin{aligned}
&e^{(n)}_{k,\mathcal{M}}|(V_{0,n}(p^2),...,V_{n,0}(p^2))\\
&= p^{(-k+\frac{1}{2}n+\frac{3}{2})(n-1)}\\
&\quad\times \left(e^{(n)}_{k,\mathcal{M}\left[X^{-1}\left(\begin{smallmatrix} p & 0\\ 0 & 1\end{smallmatrix}\right)^{-1}\right]}|U_{\left(\begin{smallmatrix} p & 0\\ 0 & 1\end{smallmatrix}\right)X\left(\begin{smallmatrix} p & 0\\ 0 & 1\end{smallmatrix}\right)},\ e^{(n)}_{k,\mathcal{M}}|U_{\left(\begin{smallmatrix} p & 0\\ 0 & 1\end{smallmatrix}\right)},\ e^{(n)}_{k,\mathcal{M}\left[\left(\begin{smallmatrix} p & 0\\ 0 & 1\end{smallmatrix}\right)\right]}\right)\\
&\quad\times \begin{pmatrix} 0 & p^{-k+1}\\ p^{-2k+2} & p^{-2k+2}\left(\frac{-m}{p}\right)\\ 0 & p^{-3k+4}\end{pmatrix} A^p_{2,n+1}(p^{k-\frac{n+2}{2}-\frac{1}{2}}).
\end{aligned}
$$

Proof. For any $\phi \in J^{(n)}_{k,\mathcal{M}}$ and for any $L = \left(\begin{smallmatrix} a & 0\\ b & 1\end{smallmatrix}\right) \in \mathbb{Z}^{(2,2)}$, a straightforward calculation gives the identity

$$
(\phi|U_L)\,|V_{\alpha,n-\alpha}(p^2) = \left(\phi|V_{\alpha,n-\alpha}(p^2)\right)|U_{\left(\begin{smallmatrix} p & 0\\ 0 & 1\end{smallmatrix}\right)^{-1}L\left(\begin{smallmatrix} p & 0\\ 0 & 1\end{smallmatrix}\right)}. \tag{7.3}
$$

We recall from Proposition 3.3 the identity

$$
\begin{aligned}
e^{(n)}_{k,\mathcal{M}} &= \sum_{d|f} g_k\left(\frac{m}{d^2}\right) E^{(n)}_{k,\mathcal{M}[W_d^{-1}]}(\tau, z\,{}^tW_d)\\
&= \sum_{d|f} g_k\left(\frac{m}{d^2}\right) E^{(n)}_{k,\mathcal{M}[W_d^{-1}]}|U_{W_d}
\end{aligned}
$$

where W_d is a matrix such that $\mathcal{M}\left[W_d^{-1}\right] \in \mathrm{Sym}_2^+$. We choose the set of matrices $\{W_d\}_d$ which satisfies $\mathcal{M}\left[W_d^{-1}\left(\begin{smallmatrix} p & 0\\ 0 & 1\end{smallmatrix}\right)^{-1}\right] \in \mathrm{Sym}_2^+$, if $d|\frac{f}{p}$. In particular, we choose $W_{pd} = \left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right) W_d$ for d such that $pd|f$. By virtue of Lemma 3.1 and of the identity $W_{pd} = \left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right) W_d$, we have

$$E^{(n)}_{k,\mathcal{M}[W_d^{-1}]}|U_{W_d\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)} \quad = \quad E^{(n)}_{k,\mathcal{M}[\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)W_{pd}^{-1}]}|U_{W_{pd}}. \tag{7.4}$$

For the sake of simplicity we write

$$\begin{aligned} E_0(d) &= E^{(n)}_{k,\mathcal{M}\left[W_{pd}^{-1}\right]}|U_{W_{pd}\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)}, \\ E_1(d) &= E^{(n)}_{k,\mathcal{M}\left[W_d^{-1}\right]}|U_{W_d\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)}, \\ E_2(d) &= E^{(n)}_{k,\mathcal{M}\left[\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)W_d^{-1}\right]}|U_{W_d}. \end{aligned} \tag{7.5}$$

We remark $E_0(d) = E_1(pd)$ and $E_1(d) = E_2(pd)$ due to the identity (7.4). From Proposition 7.4 and due to identities (7.3) and (7.5) we get

$$\begin{aligned} &\left(E^{(n)}_{k,\mathcal{M}\left[W_d^{-1}\right]}|U_{W_d}\right) | \left(V_{0,n}(p^2), \cdots, V_{n,0}(p^2)\right) \\ = \quad & p^{(-k+\frac{1}{2}n+\frac{3}{2})(n-1)} \left(E_0(d), E_1(d), E_2(d)\right) \begin{pmatrix} 0 & a_{0,\frac{m}{d^2},p,k} \\ p^{-2k+2} & a_{1,\frac{m}{d^2},p,k} \\ 0 & a_{2,\frac{m}{d^2},p,k} \end{pmatrix} \\ & \times A^p_{2,n+1}(p^{k-\frac{n+2}{2}-\frac{1}{2}}). \end{aligned}$$

Hence from Proposition 3.3 we have

$$\begin{aligned} & e^{(n)}_{k,\mathcal{M}}|(V_{0,n}(p^2), ..., V_{n,0}(p^2)) \\ = \quad & \sum_{d|f} g_k\left(\frac{m}{d^2}\right) \left(E^{(n)}_{k,\mathcal{M}\left[W_d^{-1}\right]}|U_{W_d}\right) |(V_{0,n}(p^2), ..., V_{n,0}(p^2)) \\ = \quad & p^{(-k+\frac{1}{2}n+\frac{3}{2})(n-1)} \sum_{d|f} g_k\left(\frac{m}{d^2}\right) \\ & \qquad \times \left(E_0(d), E_1(d), E_2(d)\right) \begin{pmatrix} 0 & a_{0,\frac{m}{d^2},p,k} \\ p^{-2k+2} & a_{1,\frac{m}{d^2},p,k} \\ 0 & a_{2,\frac{m}{d^2},p,k} \end{pmatrix} A^p_{2,n+1}(p^{k-\frac{n+2}{2}-\frac{1}{2}}). \end{aligned}$$

On the RHS of the above identity we obtain

$$\begin{aligned} & \sum_{d|f} g_k\left(\frac{m}{d^2}\right) \left(E_0(d), E_1(d), E_2(d)\right) \begin{pmatrix} 0 \\ p^{-2k+2} \\ 0 \end{pmatrix} \\ = \quad & p^{-2k+2} \sum_{d|f} g_k\left(\frac{m}{d^2}\right) E_1(d) \end{aligned}$$

$$
\begin{aligned}
&= p^{-2k+2} \sum_{d|f} g_k\left(\frac{m}{d^2}\right) E^{(n)}_{k,\mathcal{M}[W_d^{-1}]}|U_{W_d\left(\begin{smallmatrix} p & 0 \\ 0 & 1\end{smallmatrix}\right)} \\
&= p^{-2k+2} e^{(n)}_{k,\mathcal{M}}|U_{\left(\begin{smallmatrix} p & 0 \\ 0 & 1\end{smallmatrix}\right)}.
\end{aligned}
$$

By using Lemma 3.2 we now have

$$
\begin{aligned}
&\sum_{d|f} g_k\left(\frac{m}{d^2}\right)(E_0(d), E_1(d), E_2(d)) \begin{pmatrix} a_{0,\frac{m}{d^2},p,k} \\ a_{1,\frac{m}{d^2},p,k} \\ a_{2,\frac{m}{d^2},p,k} \end{pmatrix} \\
= &\sum_{d|f} g_k\left(\frac{m}{d^2}\right)\Bigg\{\delta\left(p|\frac{f}{d}\right)p^{-3k+4}E_0(d) \\
&+\left(p^{-2k+2}\left(\frac{-m/d^2}{p}\right)+\delta\left(p\nmid\frac{f}{d}\right)p^{-3k+4}\right)E_1(d) \\
&+\left(p^{-k+1}-p^{-2k+2}\left(\frac{-m/d^2}{p}\right)\right)E_2(d)\Bigg\} \\
= &\ p^{-3k+4}\sum_{\substack{d|f \\ \frac{f}{d}\equiv 0\ (p)}} g_k\left(\frac{m}{d^2}\right)E_0(d) + p^{-2k+2}\left(\frac{-D_0}{p}\right)\sum_{\substack{d|f \\ \frac{f}{d}\not\equiv 0\ (p)}} g_k\left(\frac{m}{d^2}\right)E_1(d) \\
&+p^{-3k+4}\sum_{\substack{d|f \\ \frac{f}{d}\not\equiv 0\ (p)}} g_k\left(\frac{m}{d^2}\right)E_1(d) \\
&+p^{-3k+4}\sum_{d|f} g_k\left(\frac{m}{d^2}\right)\left(p^{2k-3}-p^{k-2}\left(\frac{-m/d^2}{p}\right)\right)E_2(d) \\
= &\ p^{-3k+4}\delta(p|f)\sum_{d|\frac{f}{p}}\left(p^{2k-3}-p^{k-2}\left(\frac{-m/(dp)^2}{p}\right)\right)g_k\left(\frac{m}{d^2p^2}\right)E_0(d) \\
&+p^{-2k+2}\left(\frac{D_0}{p}\right)\sum_{\substack{d|f \\ \frac{f}{d}\not\equiv 0\ (p)}} g_k\left(\frac{m}{d^2}\right)E_1(d) + p^{-3k+4}\sum_{\substack{d|f \\ \frac{f}{d}\not\equiv 0\ (p)}} g_k\left(\frac{mp^2}{(dp)^2}\right)E_2(pd) \\
&+p^{-3k+4}\sum_{d|f} g_k\left(\frac{mp^2}{d^2}\right)E_2(d) \\
= &\ p^{-k+1}\delta(p|f)\sum_{d|\frac{f}{p}} g_k\left(\frac{m}{d^2p^2}\right)E_0(d) - p^{-2k+2}\delta(p|f)\left(\frac{D_0}{p}\right) \\
&\times\sum_{\substack{d|f \\ \frac{f}{d}\not\equiv 0\ (p)}} g_k\left(\frac{m}{d^2}\right)E_0\left(\frac{d}{p}\right) \\
&+p^{-2k+2}\left(\frac{D_0}{p}\right)\sum_{\substack{d|f \\ \frac{f}{d}\not\equiv 0\ (p)}} g_k\left(\frac{m}{d^2}\right)E_1(d) + p^{-3k+4}\sum_{\substack{d|fp \\ \frac{fp}{d}\not\equiv 0\ (p)}} g_k\left(\frac{mp^2}{d^2}\right)E_2(d)
\end{aligned}
$$

$$
\begin{aligned}
& +p^{-3k+4} \sum_{\substack{d|fp \\ \frac{fp}{d}\equiv 0\ (p)}} g_k\left(\frac{mp^2}{d^2}\right) E_2(d) \\
= \quad & p^{-k+1}\delta(p|f) \sum_{d|\frac{f}{p}} g_k\left(\frac{m}{d^2p^2}\right) E_0(d) + p^{-2k+2}\left(\frac{D_0 f^2}{p}\right) \sum_{d|f} g_k\left(\frac{m}{d^2}\right) E_1(d) \\
& +p^{-3k+4} \sum_{d|fp} g_k\left(\frac{mp^2}{d^2}\right) E_2(d).
\end{aligned}
$$

Hence

$$
\begin{aligned}
& \sum_{d|f} g_k\left(\frac{m}{d^2}\right) (E_0(d), E_1(d), E_2(d)) \begin{pmatrix} a_{0,\frac{m}{d^2},p,k} \\ a_{1,\frac{m}{d^2},p,k} \\ a_{2,\frac{m}{d^2},p,k} \end{pmatrix} \\
= & \left(\sum_{d|\frac{f}{p}} g_k\left(\frac{m}{d^2p^2}\right) E_0(d),\ \sum_{d|f} g_k\left(\frac{m}{d^2}\right) E_1(d),\ \sum_{d|fp} g_k\left(\frac{mp^2}{d^2}\right) E_2(d) \right) \\
& \times \begin{pmatrix} p^{-k+1} \\ p^{-2k+2}\left(\frac{-m}{p}\right) \\ p^{-3k+4} \end{pmatrix} \\
= & \left(e^{(n)}_{k,\mathcal{M}\left[X^{-1}\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)^{-1}\right]} |U_{\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)X\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)},\ e^{(n)}_{k,\mathcal{M}}|U_{\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)},\ e^{(n)}_{k,\mathcal{M}\left[\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)\right]} \right) \\
& \times \begin{pmatrix} p^{-k+1} \\ p^{-2k+2}\left(\frac{-m}{p}\right) \\ p^{-3k+4} \end{pmatrix}.
\end{aligned}
$$

Therefore

$$
\begin{aligned}
& e^{(n)}_{k,\mathcal{M}}|(V_{0,n}(p^2), ..., V_{n,0}(p^2)) \\
= \quad & p^{(-k+\frac{1}{2}n+\frac{3}{2})(n-1)} \sum_{d|f} g_k\left(\frac{m}{d^2}\right) (E_0(d),\ E_1(d),\ E_2(d)) \begin{pmatrix} 0 & a_{0,\frac{m}{d^2},p,k} \\ p^{-2k+2} & a_{1,\frac{m}{d^2},p,k} \\ 0 & a_{2,\frac{m}{d^2},p,k} \end{pmatrix} \\
& \times A^p_{2,n+1}(p^{k-\frac{n+2}{2}-\frac{1}{2}}) \\
= \quad & p^{(-k+\frac{1}{2}n+\frac{3}{2})(n-1)} \\
& \times \left(e^{(n)}_{k,\mathcal{M}\left[X^{-1}\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)^{-1}\right]} |U_{\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)X\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)},\ e^{(n)}_{k,\mathcal{M}}|U_{\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)},\ e^{(n)}_{k,\mathcal{M}\left[\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)\right]} \right) \\
& \times \begin{pmatrix} 0 & p^{-k+1} \\ p^{-2k+2} & p^{-2k+2}\left(\frac{-m}{p}\right) \\ 0 & p^{-3k+4} \end{pmatrix} A^p_{2,n+1}(p^{k-\frac{n+2}{2}-\frac{1}{2}}).
\end{aligned}
$$

□

Proposition 7.5 is a generalized Maass relation for matrix index of integral-weight. The generalized Maass relation for integer index of half-integral weight is as follows.

Theorem 7.6. *Let $e^{(n)}_{k-\frac{1}{2},m}$ be the m-th Fourier-Jacobi coefficient of generalized Cohen-Eisenstein series $H^{(n+1)}_{k-\frac{1}{2}}$. (See (1.1)). Then we obtain*

$$\begin{aligned} & e^{(n)}_{k-\frac{1}{2},m}|(\tilde{V}_{0,n}(p^2), \tilde{V}_{1,n-1}(p^2), ..., \tilde{V}_{n,0}(p^2)) \\ = \ & p^{k(n-1)-\frac{1}{2}(n^2+5n-5)} \left(e^{(n)}_{k-\frac{1}{2},\frac{m}{p^2}}|U_{p^2},\, e^{(n)}_{k-\frac{1}{2},m}|U_p,\, e^{(n)}_{k-\frac{1}{2},mp^2} \right) \\ & \times \begin{pmatrix} 0 & p^{2k-3} \\ p^{k-2} & p^{k-2}\left(\frac{-m}{p}\right) \\ 0 & 1 \end{pmatrix} A^p_{2,n+1}\left(p^{k-\frac{n+2}{2}-\frac{1}{2}}\right) diag(1, p^{1/2}, \cdots, p^{n/2}). \end{aligned}$$

Here $A^p_{2,n+1}\left(p^{k-\frac{n+2}{2}-\frac{1}{2}}\right)$ is a $2 \times (n+1)$ matrix which is introduced in the beginning of § 7 and the both side of the above identity are vectors of forms.

Proof. From Lemma 4.2 and from the definitions of $e^{(n)}_{k,\mathcal{M}}$ and $e^{(n)}_{k-\frac{1}{2},m}$, we have

$$\iota_{\mathcal{M}}(e^{(n)}_{k,\mathcal{M}}) = e^{(n)}_{k-\frac{1}{2},m}.$$

By using Proposition 4.4 we have

$$\begin{aligned} e^{(n)}_{k-\frac{1}{2},m}|\tilde{V}_{\alpha,n-\alpha}(p^2) &= \iota_{\mathcal{M}}(e^{(n)}_{k,\mathcal{M}})|\tilde{V}_{\alpha,n-\alpha}(p^2) \\ &= p^{k(2n+1)-n(n+\frac{7}{2})+\frac{1}{2}\alpha} \iota_{\mathcal{M}\left[\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)\right]}(e^{(n)}_{k,\mathcal{M}}|V_{\alpha,n-\alpha}(p^2)) \end{aligned}$$

From Proposition 4.3 we also have identities

$$\begin{aligned} e^{(n)}_{k-\frac{1}{2},\frac{m}{p^2}}|U_{p^2} &= \iota_{\mathcal{M}\left[\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)\right]} \left(e^{(n)}_{k,\mathcal{M}\left[X^{-1}\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)^{-1}\right]}|U_{\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)X\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)} \right), \\ e^{(n)}_{k-\frac{1}{2},m}|U_p &= \iota_{\mathcal{M}\left[\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)\right]} \left(e^{(n)}_{k,\mathcal{M}}|U_{\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)} \right), \end{aligned}$$

and

$$e^{(n)}_{k-\frac{1}{2},mp^2} = \iota_{\mathcal{M}\left[\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)\right]} \left(e^{(n)}_{k,\mathcal{M}\left[\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)\right]} \right).$$

Because the map $\iota_{\mathcal{M}\left[\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)\right]}$ is a linear map, this theorem follows from Proposition 7.5 and from the above identities. □

8 Maass relation for Siegel cusp forms of half-integral weight and lifts

In this section we shall prove Theorem 8.3.
We denote by $S_k(\Gamma_n) \subset M_k(\Gamma_n)$, $S^+_{k-\frac{1}{2}}(\Gamma_0^{(n)}(4)) \subset M^+_{k-\frac{1}{2}}(\Gamma_0^{(n)}(4))$, $J^{(n)\,cusp}_{k,1} \subset J^{(n)}_{k,1}$ and $J^{(n)*\,cusp}_{k-\frac{1}{2},m} \subset J^{(n)*}_{k-\frac{1}{2},m}$ the spaces of the cusp forms, respectively (cf §4.3, §4.4, §2.5 and §2.6).
Let k be an even integer and $f \in S_{2(k-n)}(\Gamma_1)$ be an eigenform for all Hecke operators. Let

$$h(\tau) \quad = \sum_{\substack{N\in\mathbb{Z} \\ N\equiv 0,3 \mod 4,\ N>0}} c(N)\, e(N\tau) \in S^+_{k-n+\frac{1}{2}}(\Gamma_0^{(1)}(4))$$

be a Hecke eigenform which corresponds to f by the Shimura correspondence. We assume that the Fourier coefficient of f at $e^{2\pi i z}$ is 1.
Let

$$I_{2n}(h)(\tau) \quad = \sum_{T\in Sym^+_{2n}} A(T)\, e(T\tau) \in S_k(\Gamma_{2n})$$

be the Ikeda lift of h. For $T \in \mathrm{Sym}^+_{2n}$ the T-th Fourier coefficient $A(T)$ of $I_{2n}(h)$ is

$$A(T) \quad = \quad c(|D_T|)\, f_T^{k-n-\frac{1}{2}} \prod_{\substack{q\,:\,prime \\ q|f_T}} \tilde{F}_q(T,\alpha_q),$$

where D_T is the fundamental discriminant and f_T is the natural number which satisfy $\det(2T) = |D_T|\, f_T^2$, and where $\left\{\alpha_q^{\pm}\right\}$ is the set of Satake parameters of f in the sense of Ikeda [Ik 01], it means that $(\alpha_q + \alpha_q^{-1})q^{k-n-1/2}$ is the q-th Fourier coefficient of f. Here $\tilde{F}_q(T,X) \in \mathbb{C}[X+X^{-1}]$ is a Laurent polynomial. For the detail of the definition of $\tilde{F}_q(T,X)$ the reader is referred to [Ik 01, page 642].
Let

$$I_{2n}(h)\left(\begin{pmatrix} \tau & z \\ {}^t z & \omega \end{pmatrix}\right) \quad = \quad \sum_{a=1}^{\infty} \psi_a(\tau,z)\, e(a\omega)$$

be the Fourier-Jacobi expansion of $I_{2n}(h)$, where $\tau \in \mathfrak{H}_{2n-1}$, $\omega \in \mathfrak{H}_1$ and $z \in \mathbb{C}^{(2n-1,1)}$. Note that $\psi_a \in J^{(2n-1)\,cusp}_{k,a}$ is a Jacobi cusp form of weight k of index a of degree $2n-1$.
By the Eichler-Zagier-Ibukiyama correspondence (see §4.3) there exists a Siegel cusp form $F \in S^+_{k-\frac{1}{2}}(\Gamma_0^{(2n-1)}(4))$ which corresponds to $\psi_1 \in J^{(2n-1)\,cusp}_{k,1}$.

For $g \in S^+_{k-1/2}(\Gamma_0^{(1)}(4))$ we put

$$\mathcal{F}_{h,g}(\tau) \quad := \quad \frac{1}{6}\int_{\Gamma_0^{(1)}(4)\backslash\mathfrak{H}_1} F\left(\begin{pmatrix}\tau & 0\\ 0 & \omega\end{pmatrix}\right)\overline{g(\omega)}\,\mathrm{Im}(\omega)^{k-\frac{5}{2}}\,d\omega$$

for $\tau \in \mathfrak{H}_{2n-2}$. It is not difficult to show that the form $\mathcal{F}_{h,g}$ belongs to $S^+_{k-\frac{1}{2}}(\Gamma_0^{(2n-2)}(4))$. The above construction of $\mathcal{F}_{h,g}$ was suggested by T.Ikeda to the author.

To show properties of $\mathcal{F}_{h,g}$ we consider the Fourier-Jacobi expansion of F. Let

$$F\left(\begin{pmatrix}\tau & z\\ {}^t z & \omega\end{pmatrix}\right) \quad = \quad \sum_{\substack{m\in\mathbb{Z}\\ m\equiv 0,3 \mod 4}} \phi_m(\tau,z)\,e(m\omega)$$

be the Fourier-Jacobi expansion of F, where $\tau \in \mathfrak{H}_{2n-2}$, $\omega \in \mathfrak{H}_1$ and $z \in \mathbb{C}^{(2n-2,1)}$. Note that $\phi_m \in J^{(2n-2)*\,cusp}_{k-\frac{1}{2},m}$ is a Jacobi cusp form of weight $k-\frac{1}{2}$ of index m and of degree $2n-2$.

Let

$$\phi_m(\tau,z) \quad = \sum_{\substack{M\in Sym^+_{2n-2},\ S\in\mathbb{Z}^{(2n-2,1)}\\ 4Mm-S^tS>0}} C_m(M,S)\,e(M\tau+S^tz)$$

be the Fourier expansion of ϕ_m, where $\tau \in \mathfrak{H}_{2n-2}$ and $z \in \mathbb{C}^{(2n-2,1)}$. We have the diagram

$$\begin{tikzcd}
 & I_{2n}(h) \in S_k(\Gamma_{2n}) \arrow[d, "\text{1st F-J}"] & \\
 & \psi_1 \in J^{(2n-1)\,cusp}_{k,1} \arrow[r, "\text{E-Z-I}"'] & F \in S^+_{k-\frac{1}{2}}(\Gamma_0^{(2n-1)}(4)) \arrow[d, "\text{F-J}"] \\
 & & \{\phi_m\}_m \in \bigotimes_{m\equiv 0,3 \bmod 4} J^{(2n-2)*\,cusp}_{k-\frac{1}{2},m} \\
h \in S^+_{k-n+\frac{1}{2}}(\Gamma_0^{(1)}(4)) \arrow[uuur, "\text{Ikeda lift}"] \arrow[r, leftrightarrow] & f \in S_{2(k-n)}(\Gamma_1). &
\end{tikzcd}$$

Lemma 8.1. *The (M,S)-th Fourier coefficient $C_m(M,S)$ of ϕ_m is*

$$C_m(M,S) \quad = \quad c(|D_T|)\,f_T^{k-n-\frac{1}{2}}\prod_{q|f_T}\tilde{F}_q(T,\alpha_q),$$

where $T \in Sym^+_{2n}$ is the matrix which satisfies

$$T = \begin{pmatrix} N & \frac{1}{2}R \\ \frac{1}{2}{}^tR & 1 \end{pmatrix}$$

and $N \in Sym^+_{2n-1}$ and $R \in \mathbb{Z}^{(2n-1,1)}$ are the matrices which satisfy

$$4N - R^tR = \begin{pmatrix} M & \frac{1}{2}S \\ \frac{1}{2}{}^tS & m \end{pmatrix}.$$

Proof. The Fourier expansion of ψ_1 is

$$\psi_1(\tau, z) = \sum_{\substack{N \in Sym^+_{2n-1},\, R \in \mathbb{Z}^{(2n-1,1)} \\ 4N - R^tR > 0}} A\left(\begin{pmatrix} N & \frac{1}{2}R \\ \frac{1}{2}{}^tR & 1 \end{pmatrix}\right) e(N\tau + R^tz).$$

And the Fourier expansion of F is

$$F(\tau) = \sum_{4N - R^tR > 0} A\left(\begin{pmatrix} N & \frac{1}{2}R \\ \frac{1}{2}{}^tR & 1 \end{pmatrix}\right) e((4N - R^tR)\tau).$$

Since ϕ_m is the m-th Fourier-Jacobi coefficient of F, the (M,S)-th Fourier coefficient $C_m(M,S)$ of ϕ_m is $A(T)$, where T is in the statement of this lemma. □

The following theorem is a generalization of the Maass relation for Siegel cusp forms of half-integral weight.

Theorem 8.2. *Let ϕ_m be the m-th Fourier-Jacobi coefficient of F as above. Then we obtain*

$$\begin{aligned} & \phi_m | (\tilde{V}_{0,2n-2}(p^2), \tilde{V}_{1,2n-3}(p^2), ..., \tilde{V}_{2n-2,0}(p^2)) \\ = \;& p^{k(2n-3)-2n^2-n+\frac{11}{2}} \left(\phi_{\frac{m}{p^2}} | U_{p^2},\, \phi_m | U_p,\, \phi_{mp^2}\right) \begin{pmatrix} 0 & p^{2k-3} \\ p^{k-2} & p^{k-2}\left(\frac{-m}{p}\right) \\ 0 & 1 \end{pmatrix} \\ & \times A^p_{2,2n-1}(\alpha_p)\; diag(1, p^{\frac{1}{2}}, p, ..., p^{n-1}) \end{aligned}$$

for any prime p, where the $2 \times (n+1)$-matrix $A^p_{2,2n-1}(\alpha_p)$ is introduced in the beginning of § 7.

Proof. Let

$$\left(\phi_m | \tilde{V}_{\alpha, 2n-2-\alpha}(p^2)\right)(\tau, z) = \sum_{\substack{M \in Sym^+_{2n-2},\; S \in \mathbb{Z}^{(2n-2,1)} \\ 4Mmp^2 - S^tS > 0}} C_m(\alpha; M, S)\; e(M\tau + S^tz).$$

be the Fourier expansion of $\phi_m|\tilde{V}_{\alpha,2n-2-\alpha}(p^2)$. We first calculate the Fourier coefficients $C_m(\alpha;M,S)$. There exist matrices $N \in \mathbb{Z}^{(2n-1,2n-1)}$ and $R \in \mathbb{Z}^{(2n-1,1)}$ which satisfy $4N - R\,{}^tR = \begin{pmatrix} M & \frac{1}{2}S \\ \frac{1}{2}{}^tS & mp^2 \end{pmatrix}$. We put $T = \begin{pmatrix} N & \frac{1}{2}R \\ \frac{1}{2}{}^tR & 1 \end{pmatrix}$. Due to Proposition 4.4 and due to the definition of $\tilde{V}_{\alpha,2n-2-\alpha}(4)$ in §4.7, we can take N and R which satisfy

$$T = \begin{pmatrix} N' & \frac{1}{2}R' \\ \frac{1}{2}{}^tR' & \mathcal{M}\left[\left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)\right] \end{pmatrix}$$

with matrices $N' \in \mathbb{Z}^{(2n-2,2n-2)}$ and $R' \in \mathbb{Z}^{(2n-2,2)}$.

We assume that p is an odd prime. Let

$$\left\{ \left(\begin{pmatrix} p^2\,{}^tD_i^{-1} & B_i \\ 0 & D_i \end{pmatrix}, \gamma_i\; p^{-n+1} \left(\det D_i\right)^{\frac{1}{2}} \right) \right\}_i$$

be a complete set of the representatives of $\Gamma_0^{(n)}(4)^* \backslash \Gamma_0^{(n)}(4)^* Y \Gamma_0^{(n)}(4)^*$, where Y is $Y = (\mathrm{diag}(1_\alpha, p1_{2n-2-\alpha}, p^2 1_\alpha, p1_{2n-2-\alpha}), p^{\alpha/2})$ and γ_i is a root of unity (see [Zh 83, Prop.7.1] or [Zh 84, Lemma 3.2] for the detail of these representatives). Then by a straightforward calculation and from Lemma 8.1 we obtain

$$\begin{aligned} C_m(\alpha;M,S) &= p^{k(2n-3)+2n^2-\frac{1}{2}-4n(n-1)} c(|D_T|) f_T^{k-n-\frac{1}{2}} \\ &\quad \times \sum_i \gamma_i \left(\det D_i\right)^{-n} e\left(\frac{1}{p^2} N\,{}^tD_i B_i\right) \prod_{q | f_{T[Q_i]}} \tilde{F}_q\left(T\left[Q_i\right], \alpha_q\right), \end{aligned} \tag{8.1}$$

where D_T is the fundamental discriminant and $f_T > 0$ is the natural number which satisfy $\det(2T) = |D_T| {f_T}^2$, and where $Q_i = \mathrm{diag}(p^{-1}\,{}^tD_i, p^{-1}, 1) \in \mathbb{Q}^{(2n,2n)}$. The number $c(|D_T|)$ is the $|D_T|$-th Fourier coefficient of h.

By virtue of the definition of $\tilde{V}_{\alpha,2n-2-\alpha}(4)$ the identity (8.1) also holds for $p = 2$.

For any prime p the (M,S)-th Fourier coefficients of $\phi_{\frac{m}{p^2}}|U_{p^2}$, $\phi_m|U_p$ and ϕ_{mp^2} are $C_{\frac{m}{p^2}}(M, p^{-2}S)$, $C_m(M, p^{-1}S)$ and $C_{mp^2}(M,S)$, respectively. These are

$$\begin{aligned} C_{\frac{m}{p^2}}(M, p^{-2}S) &= p^{-2(k-n-\frac{1}{2})} c(|D_T|) f_T^{k-n-\frac{1}{2}} \prod_{q | f_T p^{-2}} \tilde{F}_q\left(T_0, \alpha_q\right), \\ C_m(M, p^{-1}S) &= p^{-(k-n-\frac{1}{2})} c(|D_T|) f_T^{k-n-\frac{1}{2}} \prod_{q | f_T p^{-1}} \tilde{F}_q\left(T_1, \alpha_q\right) \end{aligned}$$

and

$$C_{mp^2}(M,S) = c(|D_T|) f_T^{k-n-\frac{1}{2}} \prod_{q | f_T} \tilde{F}_q\left(T, \alpha_q\right),$$

respectively, where we put $T_0 = T\left[\begin{pmatrix} 1_{2n-2} & 0 & 0 \\ 0 & p^{-2} & 0 \\ 0 & 0 & 1 \end{pmatrix}\right]$ and $T_1 = T\left[\begin{pmatrix} 1_{2n-2} & 0 & 0 \\ 0 & p^{-1} & 0 \\ 0 & 0 & 1 \end{pmatrix}\right]$. Note that if $p^{-1}S \in \mathbb{Z}^{(2n-2,1)}$, then f_T is divisible by p, and if $p^{-2}S \in \mathbb{Z}^{(2n-2,1)}$, then f_T is divisible by p^2.
Note that the Fourier coefficients of $e^{(2n-2)}_{k-\frac{1}{2},m}|\tilde{V}_{\alpha,2n-2-\alpha}(p^2)$, $e^{(2n-2)}_{k-\frac{1}{2},\frac{m}{p^2}}|U_{p^2}$, $e^{(2n-2)}_{k-\frac{1}{2},m}|U_p$ and $e^{(2n-2)}_{k-\frac{1}{2},mp^2}$ have the same form of the above expressions by substituting $\alpha_q = q^{k-n-\frac{1}{2}}$ and by replacing $c(|D_T|)$ by $c\, h_{k-n+\frac{1}{2}}(|D_T|)$, where $h_{k-n+\frac{1}{2}}(|D_T|)$ is the $|D_T|$-th Fourier coefficient of the Cohen-Eisenstein series $\mathcal{H}^{(1)}_{k-n+\frac{1}{2}}$ of weight $k-n+\frac{1}{2}$, and where $c := c_{k,2n} = 2^n\zeta(1-k)^{-1}\prod_{i=1}^n \zeta(1+2i-2k)^{-1}$. On the other hand, Theorem 7.6 is valid for infinitely many integer k. Therefore Theorem 7.6 deduces not only the relation among the Fourier coefficients of three forms $e^{(2n-2)}_{k-\frac{1}{2},\frac{m}{p^2}}$, $e^{(2n-2)}_{k-\frac{1}{2},m}$ and $e^{(2n-2)}_{k-\frac{1}{2},mp^2}$, but also the relation among the polynomials $\{\tilde{F}_q(T,X)\}_T$ of X. (cf. [Ik 01, Lemma 10.5 and page 665. line 2]. More precisely, we can conclude that the polynomial

$$p^{k(2n-3)+2n^2-\frac{1}{2}-4n(n-1)}\sum_i \gamma_i\,(\det D_i)^{-n}\, e\!\left(\frac{1}{p^2}N^t D_i B_i\right)\prod_{q|f_{T[Q_i]}}\tilde{F}_q\left(T\left[Q_i\right],X\right)$$

of X coincides with the $(\alpha+1)$-th component of the vector

$$\begin{aligned}&p^{k(2n-3)-2n^2-n+\frac{11}{2}}\\ &\times\left(p^{-2(k-n-\frac{1}{2})}\prod_{q|f_Tp^{-2}}\tilde{F}_q\left(T_0,X\right),\ p^{-(k-n-\frac{1}{2})}\prod_{q|f_Tp^{-1}}\tilde{F}_q\left(T_1,X\right),\ \prod_{q|f_T}\tilde{F}_q\left(T,X\right)\right)\\ &\times\begin{pmatrix} 0 & p^{2k-3} \\ p^{k-2} & p^{k-2}\left(\frac{-m}{p}\right) \\ 0 & 1\end{pmatrix} A^p_{2,2n-1}(X)\ \mathrm{diag}(1,p^{1/2},...,p^{(2n-2)/2}).\end{aligned}$$

Therefore $C_m(\alpha;M,S)$ coincides the $(\alpha+1)$-th component of the vector

$$\begin{aligned}&p^{k(2n-3)-2n^2-n+\frac{11}{2}}\left(C_{\frac{m}{p^2}}(M,p^{-2}S),\, C_m(M,p^{-1}S),\, C_{mp^2}(M,S)\right)\\ &\times\begin{pmatrix} 0 & p^{2k-3} \\ p^{k-2} & p^{k-2}\left(\frac{-m}{p}\right) \\ 0 & 1\end{pmatrix} A^p_{2,2n-1}(\alpha_p)\ \mathrm{diag}(1,p^{1/2},...,p^{(2n-2)/2}).\end{aligned}$$

Thus we conclude this theorem. □

Let $\tilde{T}_{\alpha,2n-2-\alpha}(p^2)$ be the Hecke operator introduced in §2.8 and let $L(s,\mathcal{F})$ be the L-function for a Hecke eigenform $\mathcal{F} \in S^+_{k-\frac{1}{2}}(\Gamma_0^{(m)}(4))$ introduced in §2.9.

THEOREM 8.3. *Let k be an even integer and n be an integer greater than 1. Let $h \in S^+_{k-n+\frac{1}{2}}(\Gamma_0^{(1)}(4))$ and $g \in S^+_{k-\frac{1}{2}}(\Gamma_0^{(1)}(4))$ be eigenforms for all Hecke operators. Then there exists a $\mathcal{F}_{h,g} \in S^+_{k-\frac{1}{2}}(\Gamma_0^{(2n-2)})$. Under the assumption that $\mathcal{F}_{h,g}$ is not identically zero, then $\mathcal{F}_{h,g}$ is an eigenform with the L-function which satisfies*

$$L(s,\mathcal{F}_{h,g}) \quad = \quad L(s,g)\prod_{i=1}^{2n-3} L(s-i,h).$$

Proof. The construction of $\mathcal{F}_{h,g}$ is stated in the above:

$$\mathcal{F}_{h,g}(\tau) \quad = \quad \frac{1}{6}\int_{\Gamma_0^{(1)}(4)\backslash\mathfrak{H}_1} F\left(\begin{pmatrix}\tau & 0\\ 0 & \omega\end{pmatrix}\right)\overline{g(\omega)}\,\mathrm{Im}(\omega)^{k-\frac{5}{2}}\,d\omega,$$

where $F \in S^+_{k-\frac{1}{2}}(\Gamma_0^{(2n-1)}(4))$ is constructed from h. By the definition of $\tilde{V}_{\alpha,2n-2-\alpha}(p^2)$ and due to Theorem 8.2 we have

$$\begin{aligned}
&\phi_m(\tau,0)|\left(\tilde{T}_{0,2n-2}(p^2),...,\tilde{T}_{2n-2,0}(p^2)\right)\\
= \;& \left(\phi_m|\left(\tilde{V}_{0,2n-2}(p^2),...,\tilde{V}_{2n-2,0}(p^2)\right)\right)(\tau,0)\\
= \;& p^{k(2n-3)-2n^2-n+\frac{11}{2}}\left(\left(\phi_{\frac{m}{p^2}}|U_{p^2}\right)(\tau,0),\ \left(\phi_m|U_p\right)(\tau,0),\ \phi_{mp^2}(\tau,0)\right)\\
& \times\begin{pmatrix}0 & p^{2k-3}\\ p^{k-2} & p^{k-2}\left(\frac{-m}{p}\right)\\ 0 & 1\end{pmatrix} A^p_{2,2n-1}(\alpha_p)\,\mathrm{diag}(1,p^{1/2},\cdots,p^{(2n-2)/2})\\
= \;& p^{k(2n-3)-2n^2-n+\frac{11}{2}}\left(\phi_{\frac{m}{p^2}}(\tau,0),\ \phi_m(\tau,0),\ \phi_{mp^2}(\tau,0)\right)\\
& \times\begin{pmatrix}0 & p^{2k-3}\\ p^{k-2} & p^{k-2}\left(\frac{-m}{p}\right)\\ 0 & 1\end{pmatrix} A^p_{2,2n-1}(\alpha_p)\,\mathrm{diag}(1,p^{1/2},\cdots,p^{(2n-2)/2}).
\end{aligned}$$

We remark

$$\begin{aligned}
&\sum_{\substack{m\\ m\equiv 0,3 \bmod 4}}\left(p^{2k-3}\phi_{\frac{m}{p^2}}(\tau,0)+p^{k-2}\left(\frac{-m}{p}\right)\phi_m(\tau,0)+\phi_{mp^2}(\tau,0)\right)e(m\omega)\\
= \;& F\left(\begin{pmatrix}\tau & 0\\ 0 & \omega\end{pmatrix}\right)|\tilde{T}_{1,0}(p^2).
\end{aligned}$$

Thus

$$
\begin{aligned}
& F\left(\begin{pmatrix}\tau & 0\\ 0 & \omega\end{pmatrix}\right)\Bigg| \left(\tilde{T}_{0,2n-2}(p^2),...,\tilde{T}_{2n-2,0}(p^2)\right)\\
= & \sum_{\substack{m\\ m\equiv 0,3 \bmod 4}} \left\{\phi_m(\tau,0)\Bigg| \left(\tilde{T}_{0,2n-2}(p^2),...,\tilde{T}_{2n-2,0}(p^2)\right)\right\} e(m\omega)\\
= & p^{k(2n-3)-2n^2-n+\frac{11}{2}} \sum_{\substack{m\\ m\equiv 0,3 \bmod 4}} \Bigg\{ \left(\phi_{\frac{m}{p^2}}(\tau,0),\ \phi_m(\tau,0),\ \phi_{mp^2}(\tau,0)\right)\\
& \times \begin{pmatrix} 0 & p^{2k-3}\\ p^{k-2} & p^{k-2}\left(\frac{-m}{p}\right)\\ 0 & 1\end{pmatrix} e(m\omega)\Bigg\}\ A^p_{2,2n-1}(\alpha_p)\,\mathrm{diag}(1,p^{1/2},\cdots,p^{(2n-2)/2})\\
= & p^{k(2n-3)-2n^2-n+\frac{11}{2}}\left(F\left(\begin{pmatrix}\tau & 0\\ 0 & \omega\end{pmatrix}\right)\Bigg|_\omega \left(\tilde{T}_{0,1}(p^2),\ \tilde{T}_{1,0}(p^2)\right)\right)\\
& \times A^p_{2,2n-1}(\alpha_p)\,\mathrm{diag}(1,p^{1/2},\cdots,p^{(2n-2)/2}).
\end{aligned}
$$

Hence

$$
\begin{aligned}
& \mathcal{F}_{h,g}| \left(\tilde{T}_{0,2n-2}(p^2),...,\tilde{T}_{2n-2,0}(p^2)\right)\\
= & \int_{\Gamma_0^{(1)}(4)\backslash\mathfrak{H}_1} \left(F\left(\begin{pmatrix}\tau & 0\\ 0 & \omega\end{pmatrix}\right)\Bigg|_\tau \left(\tilde{T}_{0,2n-2}(p^2),...,\tilde{T}_{2n-2,0}(p^2)\right)\right)\\
& \qquad \times\overline{g(\omega)}\,\mathrm{Im}(\omega)^{k-\frac{5}{2}}\,d\omega\\
= & p^{k(2n-3)-2n^2-n+\frac{11}{2}}\\
& \times \int_{\Gamma_0^{(1)}(4)\backslash\mathfrak{H}_1} \left(F\left(\begin{pmatrix}\tau & 0\\ 0 & \omega\end{pmatrix}\right)\Bigg|_\omega \left(\tilde{T}_{0,1}(p^2),\ \tilde{T}_{1,0}(p^2)\right)\right)\overline{g(\omega)}\,\mathrm{Im}(\omega)^{k-\frac{5}{2}}\,d\omega\\
& \times A^p_{2,2n-1}(\alpha_p)\,\mathrm{diag}(1,p^{1/2},\cdots,p^{(2n-2)/2}).
\end{aligned}
$$

Let $b(p)$ be the eigenvalue of g with respect to $\tilde{T}_{1,0}(p^2)$. We remark that $b(p)$ is a real number. We have

$$
\begin{aligned}
& \mathcal{F}_{h,g}| \left(\tilde{T}_{0,2n-2}(p^2),...,\tilde{T}_{2n-2,0}(p^2)\right)\\
& = p^{k(2n-3)-2n^2-n+\frac{11}{2}}\mathcal{F}_{h,g}(\tau)\\
& \quad \times\left\{(p^{k-2},b(p))A^p_{2,2n-1}(\alpha_p)\,\mathrm{diag}(1,p^{1/2},\cdots,p^{(2n-2)/2})\right\}.
\end{aligned}
\tag{8.2}
$$

Therefore $\mathcal{F}_{h,g}$ is an eigenform for any $\tilde{T}_{\alpha,2n-2-\alpha}(p^2)$.
Let $\{\beta_p^{\pm}\}$ be the set of complex numbers which satisfy

$$
1-b(p)z+p^{2k-3}z^2 \quad = \quad (1-\beta_p p^{k-3/2}z)(1-\beta_p^{-1}p^{k-3/2}z).
$$

Let $\{\mu_{0,p}^2,\mu_{1,p}^{\pm},...\mu_{2n-2,p}^{\pm}\}$ be the p-parameters of $\mathcal{F}_{h,g}$ (see §2.9 for the definition of p-parameters). We remark $\mu_{0,p}^2\mu_{1,p}\cdots\mu_{2n-2,p}=p^{2(n-1)(k-n)}$.

We now assume that p is an odd prime.
Let $\Psi_{2n-2}(K_\alpha^{(2n-2)}) \in R_{2n-2}$ be the Laurent polynomial of $\{z_i\}_{i=0,\dots,2n-2}$ introduced in §2.9. The explicit formula of $\Psi_{2n-2}(K_\alpha^{(2n-2)})$ was obtained in Proposition 6.3. The eigenvalue of $\mathcal{F}_{h,g}$ for $\tilde{T}_{\alpha,2n-2-\alpha}(p^2)$ $(\alpha = 0,\dots,2n-2)$ is obtained by substituting $z_i = \mu_i$ into $\Psi_{2n-2}(K_\alpha^{(2n-2)})$. We remark that the eigenvalue of $\mathcal{F}_{h,g}$ for $\tilde{T}_{0,2n-2}(p^2)$ is $p^{(n-1)(2k-4n+1)}$.
From the identities (8.2) and (6.5), we obtain

$$\begin{aligned}
&p^{2n^2-6n+5}(p^{-1/2}, \mu_{1,p} + \mu_{1,p}^{-1}) \prod_{l=2}^{2n-2} B_{l,l+1}(\mu_{l,p}) \\
&= p^{2n^2-6n+5}(p^{-1/2}, \beta_p + \beta_p^{-1}) \prod_{l=2}^{2n-2} B_{l,l+1}(p^{n-l}\alpha_p).
\end{aligned} \tag{8.3}$$

Here the components of the vectors in the above identity (8.3) are eigenvalues of $\mathcal{F}_{h,g}$ for

$$\tilde{T}_{0,2n-2}(p^2)^{-1}\tilde{T}_{\alpha,2n-2-\alpha}(p^2) \qquad (\alpha = 0,\dots,2n-2).$$

If we substitute $z_1 = \beta_p$ and $z_i = p^{n-i}\alpha_p$ $(i = 2,\dots,2n-2)$ into the Laurent polynomial $(\Psi_{2n-2}(K_0^{(2n-2)}))^{-1}\Psi_{2n-2}(K_\alpha^{(2n-2)})$, then due to (8.3) this value is the eigenvalue of $\mathcal{F}_{h,g}$ for $\tilde{T}_{0,2n-2}(p^2)^{-1}\tilde{T}_{\alpha,2n-2-\alpha}(p^2)$. Because R_{2n-2} is generated by $\Psi_{2n-2}(K_\alpha^{(2n-2)})$ $(\alpha = 0,\dots,2n-2)$ and $\Psi_{2n-2}(K_0^{(2n-2)})^{-1}$ and because of the fact that the p-parameters are uniquely determined up to the action of the Weyl group W_2, we therefore can take the p-parameters $\left\{\mu_{1,p}^{\pm},\cdots,\mu_{2n-2,p}^{\pm}\right\}$ of $\mathcal{F}_{h,g}$ as

$$\left\{\beta_p^{\pm}, p^{n-2}\alpha_p^{\pm}, p^{n-3}\alpha_p^{\pm},\cdots,p^{-n+2}\alpha_p^{\pm}\right\}.$$

Hence the Euler p-factor $Q_{\mathcal{F}_{h,g},p}(z)$ of $\mathcal{F}_{h,g}$ for odd prime p is

$$\begin{aligned}
Q_{\mathcal{F}_{h,g},p}(z) &= \prod_{i=1}^{2n-2}\left\{(1-\mu_{i,p}z)\left(1-\mu_{i,p}^{-1}z\right)\right\} \\
&= (1-\beta_p z)\left(1-\beta_p^{-1}z\right)\prod_{i=1}^{2n-3}\left\{\left(1-\alpha_p p^{-n+i}z\right)\left(1-\alpha_p^{-1}p^{-n+i}z\right)\right\}.
\end{aligned} \tag{8.4}$$

We now consider the case $p = 2$. The identity (8.2) is also valid for $p = 2$. Because $\tilde{\gamma}_{j,2}$ is defined in the same formula as in the case of odd primes, we also obtain the identity (8.4) for $p = 2$.
Thus we conclude

$$\begin{aligned}
L(s,\mathcal{F}_{h,g}) &= \prod_p \prod_{i=1}^{2n-2}\left\{\left(1-\mu_{i,p}p^{-s+k-\frac{3}{2}}\right)\left(1-\mu_{i,p}^{-1}p^{-s+k-\frac{3}{2}}\right)\right\}^{-1} \\
&= L(s,g)\prod_{i=1}^{2n-3} L(s-i,h).
\end{aligned}$$

□

9 Examples of non-vanishing

Lemma 9.1. *The form $\mathcal{F}_{h,g}$ in Theorem 8.3 is not identically zero, if* $(n,k) = (2,12)$, $(2,14)$, $(2,16)$, $(2,18)$, $(3,12)$, $(3,14)$, $(3,16)$, $(3,18)$, $(3,20)$, $(4,10)$, $(4,12)$, $(4,14)$, $(4,16)$, $(4,18)$, $(4,20)$, $(5,14)$, $(5,16)$, $(5,18)$, $(5,20)$, $(6,12)$, $(6,14)$, $(6,16)$, $(6,18)$ *or* $(6,20)$.

Proof. Let $h \in S^+_{k-n+\frac{1}{2}}(\Gamma_0^{(1)}(4))$, $F \in S^+_{k-\frac{1}{2}}(\Gamma_0^{(2n-1)}(4))$ and $\mathcal{F}_{h,g} \in S^+_{k-\frac{1}{2}}(\Gamma_0^{(2n-2)}(4))$ be the same symbols in §8. We have

$$F\left(\begin{pmatrix}\tau & 0\\ 0 & \omega\end{pmatrix}\right) = \sum_g \frac{1}{\langle g,g\rangle}\mathcal{F}_{h,g}(\tau)g(\omega). \qquad (9.1)$$

Here in the summation g runs over a basis of $S^+_{k-\frac{1}{2}}(\Gamma_0^{(1)}(4))$ which consists of Hecke eigenforms.
On the other hand, we have

$$F\left(\begin{pmatrix}\tau & 0\\ 0 & \omega\end{pmatrix}\right) = \sum_{M\in Sym^+_{2n-2},\, m\in Sym^+_1} K(M,m)e(N\tau)e(m\omega), \qquad (9.2)$$

where

$$K(M,m) = \sum_{\substack{S\in\mathbb{Z}^{(2n-2,1)}\\ 4Mm-S^tS>0}} C_m(M,S)$$

and where $C_m(M,S)$ is the $\begin{pmatrix}M & S\\ {}^tS & m\end{pmatrix}$-th Fourier coefficient of F. By using a computer algebraic system and Katsurada's formula for Siegel series [Ka 99], we can compute the explicit values of Fourier coefficients $C_m(M,S)$. Hence we can also compute some Fourier coefficients $K(M,m)$.
By virtue of the identities (9.1) and (9.2), we obtain

$$K(M,m) = \sum_g \frac{1}{\langle g,g\rangle}A(M;\mathcal{F}_{h,g})A(m;g),$$

where $A(M;\mathcal{F}_{h,g})$ is the M-th Fourier coefficient of $\mathcal{F}_{h,g}$ and where $A(m;g)$ is the m-th Fourier coefficient of g. Here Fourier coefficients $A(m;g)$ are calculated through the structure theorem of Kohnen plus space [Ko 80]. Therefore we can calculate some Fourier coefficients $A(M;\mathcal{F}_{h,g})$.
For example, if $(n,k) = (2,10)$, then $k-1/2 = 19/2$ and $k-n+1/2 = 17/2$. We have $\dim S^+_{19/2}(\Gamma_0^{(1)}(4)) = \dim S^+_{17/2}(\Gamma_0^{(1)}(4)) = 1$. Let $g \in S^+_{19/2}(\Gamma_0^{(1)}(4))$ and $h \in S^+_{17/2}(\Gamma_0^{(1)}(4))$ be Hecke eigenforms such that the Fourier coefficients satisfy $A(3;g) = A(1;h) = 1$. We remark that all Fourier coefficients of g

and h are real numbers. Let $K(M,m)$ be the number defined in (9.2), where $F \in S^{+}_{19/2}(\Gamma_0^{(3)}(4))$ is the Siegel modular form constructed from h. Because $\dim S^{+}_{19/2}(\Gamma_0^{(1)}(4)) = 1$, we need to check $K(M,m) \neq 0$ for a pair $(M,m) \in Sym^{+}_{2n-2} \times Sym^{+}_{1}$. We take $M = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$ and $m = 3$, then

$$\begin{aligned} &K(M,m) \\ &= C_3\left(M, \begin{pmatrix} 2 \\ 2 \end{pmatrix}\right) + C_3\left(M, \begin{pmatrix} 2 \\ -2 \end{pmatrix}\right) + C_3\left(M, \begin{pmatrix} -2 \\ 2 \end{pmatrix}\right) + C_3\left(M, \begin{pmatrix} -2 \\ -2 \end{pmatrix}\right) \\ &= -336 - 168 - 168 - 336 \\ &\neq 0. \end{aligned}$$

Therefore $\mathcal{F}_{h,g} \not\equiv 0$ for $(n,k) = (2,10)$.
Similarly, by using a computer algebraic system, we can also check $\mathcal{F}_{h,g} \not\equiv 0$ for any h and g for other (n,k) in the lemma. □

References

[Ar 98] T. Arakawa: Köcher-Maass Dirichlet Series Corresponding to Jacobi forms and Cohen Eisenstein Series, *Comment. Math. Univ. St. Paul.* 47 No.1 (1998), 93–122.

[Bo 83] S. Böcherer: Über die Fourier-Jacobi-Entwicklung Siegelscher Eisensteinreihen, *Math. Z.* 183 (1983), 21–46

[Co 75] H. Cohen: Sums involving the values at negative integers of L-functions of quadratic characters, *Math. Ann.* 217 (1975), 171–185.

[E-Z 85] M. Eichler and D. Zagier: Theory of Jacobi Forms, Progress in Math. 55, Birkhäuser, Boston-Basel-Stuttgart, (1985).

[H-I 05] S. Hayashida and T. Ibukiyama: Siegel modular forms of half integral weights and a lifting conjecture, *Journal of Kyoto Univ,* 45 No.3 (2005), 489–530.

[H 13] S. Hayashida: On generalized Maass relations and their application to Miyawaki-Ikeda lifts, *Comment. Math. Univ. St. Pauli*, 62 No.1 (2013), 59–90.

[Ib 92] T. Ibukiyama: On Jacobi forms and Siegel modular forms of half integral weights, *Comment. Math. Univ. St. Paul.* 41 No.2 (1992), 109–124.

[Ik 01] T. Ikeda: On the lifting of elliptic cusp forms to Siegel cusp forms of degree 2n, *Ann. of Math. (2)* 154 no.3, (2001), 641–681.

[Ik 06] T. Ikeda: Pullback of the lifting of elliptic cusp forms and Miyawaki's conjecture, *Duke Math. J.* 131, No.3, (2006), 469–497.

[Ka 99] H. Katsurada: An explicit formula for Siegel series. *Amer. J. Math.* 121, No. 2, (1999), 415–452.

[Ko 80] W. Kohnen: Modular forms of half integral weight on $\Gamma_0(4)$, *Math, Ann.* 248 (1980), 249–266.

[Ko 02] W. Kohnen: Lifting modular forms of half-integral weight to Siegel modular forms of even genus, *Math, Ann.*, 322 (2002), 787–809.

[KK 05] W. Kohnen and H. Kojima: A Maass space in higher genus, *Compos. Math.*, 141, No. 2, (2005), 313–322.

[Kr 86] A. Krieg: Das Vertauschungsgesetz zwischen Hecke-Operatoren und dem Siegelschen ϕ-Opera-tor, *Arch. Math.* 46, No. 4, (1986), 323–329.

[OKK 89] Y.-Y. Oh, J.-K. Koo and M.-H. Kim: Hecke operators and the Siegel operator, *J. Korean Math. Soc.* 26, No. 2, (1989), 323–334.

[Sh 73] G. Shimura: On modular forms of half-integral weight, *Ann. of Math.* 97, (1973), 440–481.

[Ta 86] Y. Tanigawa: Modular descent of Siegel modular forms of half integral weight and an analogy of the Maass relation, *Nagoya Math. J.* 102 (1986), 51–77.

[Yn 10] S. Yamana: Maass relations in higher genus, *Math. Z.*, 265, no. 2, (2010), 263–276.

[Yk 86] T. Yamazaki: Jacobi forms and a Maass relation for Eisenstein series, *J. Fac. Sci. Univ. Tokyo Sect. IA, Math.* 33 (1986), 295–310.

[Yk 89] T. Yamazaki: Jacobi forms and a Maass relation for Eisenstein series II, *J. Fac. Sci. Univ. Tokyo Sect. IA, Math.* 36 (1989), 373–386.

[Zh 83] V. G. Zhuravlev: Hecke rings for a covering of the symplectic group, *Mat. Sb.* 121 (163) (1983), 381–402.

[Zh 84] V. G. Zhuravlev: Euler expansions of theta transforms of Siegel modular forms of half-integral weight and their analytic properties, *Mat. Sb.* 123 (165) (1984), 174–194.

[Zi 89] C. Ziegler: Jacobi forms of higher degree, *Abh. Math. Sem. Univ. Hamburg.* 59 (1989), 191–224.

Shuichi Hayashida
Department of Mathematics
Joetsu University of
Education
1 Yamayashikimachi
Joetsu, Niigata 943-8512
JAPAN
e-mail hayasida@juen.ac.jp

On Asymptotic Bounds for the Number of Irreducible Components of the Moduli Space of Surfaces of General Type II

Michael Lönne and Matteo Penegini

Received: July 29, 2015
Revised: December 12, 2015

Communicated by Thomas Peternell

Abstract. In this paper we investigate the asymptotic growth of the number of irreducible and connected components of the moduli space of surfaces of general type corresponding to certain families of surfaces isogenous to a higher product with group $(\mathbb{Z}/2\mathbb{Z})^k$. We obtain a significantly higher growth than the one in our previous paper [LP14].

2010 Mathematics Subject Classification: 14J10, 14J29, 20D15, 20D25, 20H10, 30F99.
Keywords and Phrases: moduli spaces, surfaces of general type, Hurwitz action, surfaces isogenous to a product

1 Introduction

It is well known (see [Gie77]) that once two positive integers x, y are fixed there exists a quasi-projective coarse moduli space $\mathcal{M}_{y,x}$ of canonical models of surfaces of general type with $x = \chi(S) = \chi(\mathcal{O}_S)$ and $y = K_S^2$. The number $\iota(x, y)$, resp. $\gamma(x, y)$, of irreducible, resp. connected, components of $\mathcal{M}_{y,x}$ is bounded from above by a function of y. In fact, Catanese proved that the number $\iota^0(x, y)$ of components containing regular surfaces, i.e., $q(S) = 0$, has an exponential upper bound in K^2. More precisely [Cat92, p.592] gives the following inequality

$$\iota^0(x, y) \leq y^{77y^2}.$$

This result is not known to be sharp and in recent papers [M97, Ch96, GP14, LP14] inequalities are proved which tell how close one can get to this bound from below. In particular, in the last two papers the authors considered families of surfaces isogenous to a product in order to construct many irreducible components of the moduli space of surfaces of general type. The reason why one works with these surfaces, is the fact that the number of families of these surfaces can be easily computed using group theoretical and combinatorial methods.

In our previous work [LP14] we constructed many such families with many different $2-$groups. There, we exploited the fact that the number of $2-$groups with given order grows very fast in function of the order. In this paper we obtain a significantly better lower bound for $\iota^0(x,y)$ using only the groups $(\mathbb{Z}/2\mathbb{Z})^k$ and again some properties of the moduli space of surfaces isogenous to a product. Our main result is the following theorem.

Theorem 1.1. *Let $h = h_{k,l}$ be number of connected components of the moduli space of surfaces of general type which contain regular surfaces isogenous to a product of holomorphic Euler number $x_{k,l} = 2^{l-3}(k^2+k-4)$ given by a $(\mathbb{Z}/2\mathbb{Z})^k$ action with ramification structure of type $(2^{k(k+1)}, 2^{4+2^{l-k+1}})$.*
If k,l are positive integers with $l > 2k$, then

$$h \quad \geq \quad 2^{\sqrt[2+\nu]{x_k}} \qquad \textit{for} \quad k \to \infty,$$

where ν is the positive real number such that $l = (2+\nu)k$. In particular, given any sequence α_i which is positive, increasing and bounded by $\frac{1}{2}$ from above, we obtain increasing sequences x_i and $y_i = 8x_i$ with

$$\iota^0(x_i, y_i) \quad \geq \quad y_i^{(y_i^{\alpha_i})}.$$

Let us explain now the way in which this paper is organized.

In the next section *Preliminaries* we recall the definition and the properties of surfaces isogenous to a higher product and the its associated group theoretical data. Moreover, we recall a result of Bauer–Catanese [BC] which allows us to count the number of connected components of the moduli space of surfaces isogenous to a product with given group and type of ramification structure.

In the last section we give the proof of the Theorem 1.1.

Acknowledgement The first author was supported by the ERC 2013 Advanced Research Grant - 340258 - TADMICAMT, at the Universität Bayreuth. The second author acknowledges Riemann Fellowship Program of the Leibniz Universität Hannover. We thank the anonymous referee for reading very carefully the manuscript and for many questions and suggestions that helped us to improve the paper a lot.

Notation and conventions. We work over the field $\mathbb{C}$ of complex numbers. By *surface* we mean a projective, non-singular surface S. For such a surface $\omega_S = \mathcal{O}_S(K_S)$ denotes the canonical bundle, $p_g(S) = h^0(S, \omega_S)$ is the *geometric genus*, $q(S) = h^1(S, \omega_S)$ is the *irregularity*, $\chi(\mathcal{O}_S) = \chi(S) = 1 - q(S) + p_g(S)$ is the *Euler-Poincaré characteristic* and $e(S)$ is the *topological Euler number* of S.

2 Preliminaries

Definition 2.1. *A surface S is said to be* isogenous to a higher product of curves *if and only if, S is a quotient $(C_1 \times C_2)/G$, where C_1 and C_2 are curves of genus at least two, and G is a finite group acting freely on $C_1 \times C_2$.*

Using the same notation as in Definition 2.1, let S be a surface isogenous to a product, and $G^\circ := G \cap (Aut(C_1) \times Aut(C_2))$. Then G° acts on the two factors C_1, C_2 and diagonally on the product $C_1 \times C_2$. If G° acts faithfully on both curves, we say that $S = (C_1 \times C_2)/G$ is a *minimal realization*. In [Cat00] it is also proven that any surface isogenous to a product admits a unique minimal realization.

Assumptions. In the following we always assume:

1. Any surface S isogenous to a product is given by its unique minimal realization;

2. $G^\circ = G$, this case is also known as *unmixed type*, see [Cat00].

Under these assumption we have.

Proposition 2.2. *[Cat00] Let $S = (C_1 \times C_2)/G$ be a surface isogenous to a higher product of curves, then S is a minimal surface of general type with the following invariants:*

$$\chi(S) = \frac{(g(C_1) - 1)(g(C_2) - 1)}{|G|}, \quad e(S) = 4\chi(S), \quad K_S^2 = 8\chi(S). \tag{1}$$

The irregularity of these surfaces is computed by

$$q(S) = g(C_1/G) + g(C_2/G). \tag{2}$$

Among the nice features of surfaces isogenous to a product, one is that their deformation class can be obtained in a purely algebraic way. Let us briefly recall this in the particular case when S is regular, i.e., $q(S) = 0$, $C_i/G \cong \mathbb{P}^1$.

Definition 2.3. *Let G be a finite group and $r \in \mathbb{N}$ with $r \geq 2$.*

- *An $r-$tuple $T = (v_1, \ldots, v_r)$ of elements of G is called a* spherical system of generators *of G if $\langle v_1, \ldots, v_r \rangle = G$ and $v_1 \cdot \ldots \cdot v_r = 1$.*

- *We say that T has an* unordered type $\tau := (m_1, \ldots, m_r)$ *if the orders of $(v_1, \ldots, v_r)$ are $(m_1, \ldots, m_r)$ up to a permutation, namely, if there is a permutation $\pi \in \mathfrak{S}_r$ such that*

$$\mathrm{ord}(v_1) = m_{\pi(1)}, \ldots, \mathrm{ord}(v_r) = m_{\pi(r)}.$$

- *Moreover, two spherical systems* $T_1 = (v_{1,1}, \ldots, v_{1,r_1})$ *and* $T_2 = (v_{2,1}, \ldots, v_{2,r_2})$ *are said to be* disjoint*, if:*

$$\Sigma(T_1) \bigcap \Sigma(T_2) = \{1\}, \tag{3}$$

where

$$\Sigma(T_i) := \bigcup_{g \in G} \bigcup_{j=0}^{\infty} \bigcup_{k=1}^{r_i} g \cdot v_{i,k}^j \cdot g^{-1}.$$

We shall also use the shorthand, for example $(2^4, 3^2)$, to indicate the tuple $(2, 2, 2, 2, 3, 3)$.

Definition 2.4. *Let* $2 < r_i \in \mathbb{N}$ *for* $i = 1, 2$ *and* $\tau_i = (m_{i,1}, \ldots, m_{i,r_i})$ *be two sequences of natural numbers such that* $m_{k,i} \geq 2$. *A* (spherical-) ramification structure *of type* (τ_1, τ_2) *and size* (r_1, r_2) *for a finite group* G*, is a pair* (T_1, T_2) *of disjoint spherical systems of generators of* G*, whose types are* τ_i*, such that:*

$$\mathbb{Z} \ni \frac{|G|(-2 + \sum_{l=1}^{r_i}(1 - \frac{1}{m_{i,l}}))}{2} + 1 \geq 2, \qquad \text{for } i = 1, 2. \tag{4}$$

Remark 2.5. Following e.g., the discussion in [LP14, Section 2] we obtain that the datum of the deformation class of a regular surface S isogenous to a higher product of curves of unmixed type together with its minimal realization $S = (C_1 \times C_2)/G$ is determined by the datum of a finite group G together with two disjoint spherical systems of generators T_1 and T_2 (for more details see also [BCG06]).

Remark 2.6. Recall that from Riemann Existence Theorem a finite group G acts as a group of automorphisms of some curve C of genus g such that $C/G \cong \mathbb{P}^1$ if and only if there exist integers $m_r \geq m_{r-1} \geq \cdots \geq m_1 \geq 2$ such that G has a spherical system of generators of type $(m_1, \ldots, m_r)$ and the following Riemann-Hurwitz relation holds:

$$2g - 2 = |G|(-2 + \sum_{i=1}^{r}(1 - \frac{1}{m_i})). \tag{5}$$

Remark 2.7. Note that a group G and a ramification structure determine the main numerical invariants of the surface S. Indeed, by (1) and (5) we obtain:

$$\begin{aligned} 4\chi(S) &= |G| \cdot \left(-2 + \sum_{k=1}^{r_1}(1 - \frac{1}{m_{1,k}})\right) \cdot \left(-2 + \sum_{k=1}^{r_2}(1 - \frac{1}{m_{2,k}})\right) \\ &=: 4\chi(|G|, (\tau_1, \tau_2)). \end{aligned} \tag{6}$$

Let S be a surface isogenous to a product of unmixed type with group G and a pair of two disjoint spherical systems of generators of types (τ_1, τ_2). By (6) we have $\chi(S) = \chi(G, (\tau_1, \tau_2))$, and hence, by (1), $K_S^2 = K^2(G, (\tau_1, \tau_2)) = 8\chi(S)$.

Let us fix a group G and a pair of unmixed ramification types (τ_1, τ_2), and denote by $\mathcal{M}_{(G,(\tau_1,\tau_2))}$ the moduli space of isomorphism classes of surfaces isogenous to a product admitting these data, by [Cat00, Cat03] the space $\mathcal{M}_{(G,(\tau_1,\tau_2))}$ consists of a finite number of connected components. Indeed, there is a group theoretical procedure to count these components. In case G is abelian it is described in [BC].

Theorem 2.8. *[BC, Theorem 1.3] . Let S be a surface isogenous to a higher product of unmixed type and with $q = 0$. Then to S we attach its finite group G (up to isomorphism) and the equivalence classes of an unordered pair of disjoint spherical systems of generators (T_1, T_2) of G, under the equivalence relation generated by:*

i. Hurwitz equivalence for T_1;

ii. Hurwitz equivalence for T_2;

iii. Simultaneous conjugation for T_1 and T_2, i.e., for $\phi \in \mathrm{Aut}(G)$ we let $\big(T_1 := (v_{1,1}, \dots, v_{1,r_1}), \quad T_2 := (v_{2,1}, \dots, v_{2,r_2})\big)$ be equivalent to

$$\big(\phi(T_1) := (\phi(v_{1,1}), \dots, \phi(v_{1,r_1})), \quad \phi(T_2) := (\phi(v_{2,1}), \dots, \phi(v_{2,r_2}))\big).$$

Then two surfaces S, S' are deformation equivalent if and only if the corresponding equivalence classes of pairs of spherical generating systems of G are the same.

The Hurwitz equivalence is defined precisely in e.g., [P13]. In the cases that we will treat the Hurwitz equivalence is given only by the braid group action on T_i defined as follows. Recall the Artin presentation of the Braid group of r_1 strands

$$\mathbf{B}_{r_1} := \langle \gamma_1, \dots, \gamma_{r_1-1} | \, \gamma_i\gamma_j = \gamma_j\gamma_i \text{ for } |i-j| \geq 2, \, \gamma_{i+1}\gamma_i\gamma_{i+1} = \gamma_i\gamma_{i+1}\gamma_i \rangle.$$

For $\gamma_i \in \mathbf{B}_{r_1}$ then:

$$\gamma_i(T_1) = \gamma_i(v_1, \dots, v_{r_1}) = (v_1, \dots, v_{i+1}, v_{i+1}^{-1}v_iv_{i+1}, \dots, v_{r_1}).$$

Moreover, notice that, since we deal here with abelian groups only, the braid group action is indeed only by permutation of the elements on the spherical system of generators.

Once we fix a finite abelian group G and a pair of types (τ_1, τ_2) (of size (r_1,r_2)) of an unmixed ramification structure for G, counting the number of connected components of $\mathcal{M}_{(G,(\tau_1,\tau_2))}$ is then equivalent to the group theoretical problem of counting the number of classes of pairs of spherical systems of generators of G of type (τ_1, τ_2) under the equivalence relation given by the action of $\mathbf{B}_{r_1} \times \mathbf{B}_{r_2} \times \mathrm{Aut}(G)$, given by:

$$(\gamma_1, \gamma_2, \phi) \cdot (T_1, T_2) := \big(\phi(\gamma_1(T_1)), \phi(\gamma_2(T_2))\big), \tag{7}$$

where $\gamma_1 \in \mathbf{B}_{r_1}$, $\gamma_2 \in \mathbf{B}_{r_1}$ and $\phi \in \mathrm{Aut}(G)$, see for more details e.g., [P13].

3 Proof of Theorem 1.1

Let us consider the group $G := (\mathbb{Z}/2\mathbb{Z})^k$, with $k >> 0$ and an integer l. We want to give to G many classes of ramification structures of size $(r_1, r_2) = (k(k+1), 2^{l-k+1}+4)$. Since the elements of G have only order two we will produce in the end ramification structure of type $((2^{r_1}), (2^{r_2}))$.
Let $T_1 := (v_1, v_2, \ldots, v_{k(k+1)})$ with the following elements v_i of G

$$\begin{aligned} v_1 &= (1,0,0,\ldots,0) \\ v_2 &= (1,0,0,\ldots,0) \\ v_3 &= (0,1,0,\ldots,0) \\ v_4 &= (0,1,0,\ldots,0) \\ v_5 &= (0,1,0,\ldots,0) \\ v_6 &= (0,1,0,\ldots,0) \\ &\vdots \\ v_{k(k+1)} &= (0,0,\ldots,0,1) \end{aligned}$$

By construction the product of the elements in T_1 is 1_G and $< T_1 > \cong G$. Define the set $M := G \setminus \{0, v_1, \ldots, v_{k(k+1)}\}$. Since M has $\#M = 2^k - k - 1$ elements we can choose a bijection

$$\varphi : \quad \{n \in \mathbb{N} | n \leq 2^k - k - 1\} \longrightarrow M.$$

Let B the set of $(2^k - k - 1)$-tuples $(n_1, \ldots, n_{\#M})$ of positive integers of sum $n_1 + n_2 + \cdots + n_{\#M}$ equal to $2^{l-k} + 2$. For every element in B we get a $2^{l-k+1} + 4$-tuple T_2 of elements of G by the map

$$(n_1, \ldots, n_{2^k-k-1}) \quad \mapsto \quad T_2$$

where

$$T_2 \quad = \quad \big(\underbrace{\varphi(1), \ldots, \varphi(1)}_{2n_1}, \ldots, \underbrace{\varphi(2^k-k-1), \ldots, \varphi(2^k-k-1)}_{2n_{2^k-k-1}}\big).$$

By construction again $< T_2 > \cong G$ (for $k > 2$) and the product of the elements in T_2 is 1_G. Hence, T_2 is a spherical system of generators for G of size $2^{l-k+1}+4$. Since $G = \mathbb{Z}/2^k$ is abelian with all non-trivial elements of order two, the pair (T_1, T_2) is a ramification structure for G of the desired type, for any element in B.

Now we count how many inequivalent ramification structures of this kind we have under the action of the group defined in Theorem 2.8 and Equation (7). First, by construction, the only element in $\mathrm{Aut}(G)$, which stabilises T_1 is the identity. Next, accordingly, (T_1, T_2') and (T_1, T_2'') are equivalent if and only

if T_2' and T_2'' are in the same braid orbit. So finally, by construction, we get inequivalent pairs associated to different elements of B.
Hence the number of inequivalent such ramification structures is equal to the number of (2^k-k-1)-tuples in B of positive integers whose sum is $2^{l-k}+2$ $= r_2/2$.
Any element in B corresponds uniquely to the sequence of 2^k-k-1 integers $n_1, n_1+n_2, \dots, n_1+\dots+n_{\#M} = r_2/2$. The sequence is strongly increasing by $n_i > 0$, so the elements in B correspond bijectively to the number of choices of 2^k-k-2 integers (all but the last integer in the sequence) in the range from 1 to $r_2/2-1$, cf [F68, Section II.5, p.38]. Hence,

$$\#B \quad = \quad \binom{\frac{r_2}{2}-1}{2^k-k-2} \quad = \quad \binom{2^{l-k}+1}{2^k-k-2},$$

Let $\nu > 0$ be a rational number and let us suppose that $l = (\nu+2)\cdot k$, then we exploit Stirling's approximation of the binomial coefficient - more exactly a corresponding lower bound. To bound the binomial coefficient $\binom{m}{n}$ we use the lower bound $(m-n)^n < m!/(m-n)!$ and an upper bound on $n!$ related to Stirling's formula. In fact, cf. [F68, eqn. (9.6), p.53],

$$n! \quad = \quad e^{d_n} n^{n+\frac{1}{2}} e^{-n} \qquad \text{for} \quad d_n \quad = \quad \ln n! - (n+\frac{1}{2})\ln n + n$$

and we get an upper bound on replacing d_n by $d_1 = 1$ due to the observation [F68, eqn. (9.9), p.53] that (d_n) is decreasing. Thus we obtain

$$\begin{aligned}\binom{2^{l-k}+1}{2^k-k-2} \quad &> \quad \frac{(\frac{2^{l-k}+1}{2^k-k-2}-1)^{2^k-k-2} e^{2^k-k-2}}{e\sqrt{(2^k-k-2)}} \\ &> \quad (2^{\nu k})^{(2^k-k-2)} \cdot \frac{e^{2^k-k-2}}{e\sqrt{(2^k-k-2)}} \quad > \quad 2^{\nu k (2^k)}. \end{aligned} \tag{8}$$

Now $y_k = 8\chi(S) = 2|G|(-2+\frac{1}{2}r_1)(-2+\frac{1}{2}r_2)$ by (6) implies $y_k = 2^k \cdot 2^{l-k} \cdot (k^2+k-4) = 2^{(\nu+2)k}(k^2+k-4)$ according to $l = (2+\nu)k$. Hence we have

$$(y_k)^{\frac{1}{\nu+2}} \cdot \frac{k}{(k^2+k-4)^{\frac{1}{\nu+2}}} \quad = \quad k2^k.$$

Using this, we obtain for k large enough in the second inequality

$$h > 2^{\nu (y_k)^{\frac{1}{\nu+2}} \cdot \frac{k}{(k^2+k-4)^{\frac{1}{\nu+2}}}} \quad > \quad 2^{(y_k^{\frac{1}{\nu+2}})}. \tag{9}$$

We can bound further for k large enough

$$2^{(y_k^{\frac{1}{\nu+2}})} \quad > \quad 2^{(y_k^{\frac{1}{2\nu+2}})\frac{\ln y_k}{\ln 2}}. \tag{10}$$

We use the identity $x^{f(x)} = e^{f(x)\ln x} = 2^{f(x)\frac{1}{\ln 2}\ln x}$ to get for all $\alpha < \frac{1}{2}$

$$h > y_k^{(y_k)^\alpha}$$

if k is large enough, depending on α. □

References

[BC] I. Bauer, F. Catanese, *Some new surfaces with $p_g = q = 0$*. Proceeding of the Fano Conference. Torino (2002), 123–142.

[BCG06] I. Bauer, F. Catanese, F. Grunewald, *Chebycheff and Belyi polynomials, dessins d'enfants, Beauville surfaces and group theory.* Mediterr. J. Math. 3, (2006), 121–146.

[Cat92] F. Catanese, *Chow varieties, Hilbert schemes and moduli spaces of surfaces of general type.* J. Algebraic Geom. 1 (1992), 561–595.

[Cat00] F. Catanese, *Fibred surfaces, varieties isogenous to a product and related moduli spaces.* Amer. J. Math. 122, (2000), 1–44.

[Cat03] F, Catanese, *Moduli spaces of surfaces and real structures.* Ann. of Math. 158, (2003), 577–592.

[CLP11] F. Catanese, M. Lönne, F. Perroni, *Irreducibility of the space of dihedral covers of the projective line of a given numerical type*, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei, 22, (2011), 291–309.

[Ch96] M. Chang, *The number of components of Hilbert schemes.* Internat. J. Math. 7 (1996), 301–306.

[F68] W. Feller, *An Introduction to Probability Theory and Its Applications*, Wiley, Vol I, third edition (1968).

[GP14] S. Garion, M. Penegini, *Beauville surfaces, moduli spaces and finite groups* Comm. in Algebra 42, (2014), 2126–2155.

[Gie77] D. Gieseker, *Global moduli for surfaces of general type.* Invent. Math 43 no. 3, (1977), 233–282.

[LP14] M. Lönne, M. Penegini, *On asymptotic bounds for the number of irreducible components of the moduli space of surfaces of general type.* Rend. Circ. Mat. Palermo (2) 64 no. 3, (2015), 483–492.

[M97] M. Manetti, *Iterated double covers and connected components of moduli spaces.* Topology 36, (1997), 745–764.

[P13] M. Penegini, *Surfaces isogenous to a product of curves, braid groups and mapping class groups* in "Beauville Surfaces and Groups", Springer Proceedings in Math. and Stats., (2015), 129–148.

Michael Lönne
Mathematatik VIII
Universität Bayreuth
Universitätsstrasse 30
D-95447 Bayreuth, Germany
michael.loenne@uni-bayreuth.de

Matteo Penegini
Dipartimento di Matematica
Università degli Studi di Genova
Via Dodecaneso 35
I-16146 Genova, Italy
penegini@dima-unige.it

Cohomologie des foncteurs polynomiaux sur les groupes libres

Aurélien Djament, Teimuraz Pirashvili [1] et Christine Vespa [2]

Received: August 28, 2015
Revised: December 08, 2015

Communicated by Stefan Schwede

Resumé. We show that extension groups between two polynomial functors on free groups are the same in the category of all functors and in a subcategory of polynomial functors of bounded degree. The proof relies on functorial properties of the group ring of free groups and its filtration by powers of the augmentation ideal. We give some applications, in particular in term of homological dimension.

Résumé. On montre que les groupes d'extensions entre foncteurs polynomiaux sur les groupes libres sont les mêmes dans la catégorie de tous les foncteurs et dans une sous-catégorie de foncteurs polynomiaux de degré borné. La démonstration repose sur les propriétés fonctorielles de l'anneau de groupe des groupes libres et de sa filtration par les puissances de l'idéal d'augmentation. On donne quelques applications, notamment en termes de dimension homologique.

2010 Mathematics Subject Classification: 18A25, 18G15, 20J15 (18E15, 18G10)
Keywords and Phrases: catégories de foncteurs ; groupes libres ; foncteurs polynomiaux ; groupes d'extensions.

1. Cet auteur est partiellement soutenu par la bourse ST08/3-387 de la GNSF.
2. Cet auteur est partiellement soutenu par le projet ANR-11-BS01-0002 HOGT : Homotopie, Opérades et Groupes de Grothendieck-Teichmüller.

Introduction

Ce travail est une contribution à l'étude des groupes d'extensions dans la catégorie des foncteurs des groupes libres (de rang fini) vers les groupes abéliens. Cette catégorie a d'abord été considérée, du point de vue cohomologique, dans l'article de Jibladze et du deuxième auteur [9] (§ 5A) sur la cohomologie des théories algébriques. Les premier et troisième auteurs l'ont également utilisée, dans [4], pour établir des résultats d'annulation d'homologie stable des groupes d'automorphismes des groupes libres à coefficients tordus. Récemment, le premier auteur a montré [2] que l'algèbre homologique dans cette catégorie de foncteurs gouverne le calcul d'autres groupes d'homologie stable des groupes d'automorphismes des groupes libres à coefficients tordus. Cette algèbre homologique s'avère plus facile d'accès que dans les catégories de foncteurs entre espaces vectoriels sur un corps fini, très étudiées en raison de leurs liens avec la topologie algébrique ou la K-théorie algébrique (cf. par exemple [6]), notamment parce que le foncteur d'abélianisation possède une résolution projective explicite simple (déjà utilisée dans [9]). L'article [4] se sert également de façon cruciale de la structure des sous-catégories de *foncteurs polynomiaux* dans cette catégorie de foncteurs.

Dans le présent travail, on montre que les groupes d'extensions entre deux foncteurs polynomiaux sur les groupes libres sont les mêmes dans la catégorie de tous les foncteurs ou dans une sous-catégorie de foncteurs polynomiaux de degré donné. Un résultat similaire vaut pour les foncteurs sur la catégorie Γ des ensembles finis pointés, comme cela résulte du théorème de type Dold-Kan établi par le deuxième auteur dans [14]. En revanche, notre résultat contraste avec la situation, plus délicate, des foncteurs sur une catégorie additive (voir [13] et [3]). Nous donnons également quelques applications.

Description des résultats Commençons par quelques notations générales. Soit k un anneau, on note k-**Mod** la catégorie des k-modules à gauche. Si $\mathcal{C}$ est une petite catégorie, on note $\mathcal{F}(\mathcal{C}; k)$ la catégorie des foncteurs de $\mathcal{C}$ vers k-**Mod** (on notera simplement $\mathcal{F}(\mathcal{C})$ pour $\mathcal{F}(\mathcal{C}; \mathbb{Z})$). Si la catégorie $\mathcal{C}$ possède un objet nul et des coproduits finis, on dispose d'une notion d'*effets croisés* et de *foncteurs polynomiaux* dans $\mathcal{F}(\mathcal{C}; k)$ (voir la section 1 ci-après pour davantage de détails). L'origine de ces notions remonte à Eilenberg-MacLane ([5], *Chapter* II), lorsque $\mathcal{C}$ est une catégorie de modules, et se généralise sans difficulté au cas qu'on vient de mentionner (voir [8], § 2). Pour tout entier d, on note $\mathcal{F}_d(\mathcal{C}; k)$ la sous-catégorie pleine de $\mathcal{F}(\mathcal{C}; k)$ des foncteurs polynomiaux de degré au plus d.

On note **gr** la catégorie des groupes libres de rang fini, ou plus exactement le squelette constitué des groupes $\mathbb{Z}^{*n}$ (l'étoile désignant le produit libre). Comme signalé plus haut, notre résultat principal est le suivant :

Théorème 1. *Soient k un anneau, $d \in \mathbb{N}$ et F, G des objets de $\mathcal{F}_d(\mathbf{gr}; k)$.*

L'application linéaire graduée naturelle

$$\mathrm{Ext}^*_{\mathcal{F}_d(\mathbf{gr};k)}(F,G) \to \mathrm{Ext}^*_{\mathcal{F}(\mathbf{gr};k)}(F,G)$$

qu'induit le foncteur d'inclusion $\mathcal{F}_d(\mathbf{gr};k) \to \mathcal{F}(\mathbf{gr};k)$ *est un isomorphisme.*

Pour des catégories sources autres que **gr**, l'énoncé similaire au théorème précédent n'est généralement pas vrai. Ainsi, rappelons quelques phénomènes connus sur une petite catégorie *additive* $\mathcal{C}$. Si l'anneau k contient le corps $\mathbb{Q}$ des nombres rationnels, le foncteur d'inclusion $\mathcal{F}_d(\mathcal{C};k) \to \mathcal{F}(\mathcal{C};k)$ induit des isomorphismes entre groupes d'extensions. Ce résultat, qui fait partie du folklore, figure dans [3] (théorème 1.2). En général, même sous de bonnes hypothèses sur la catégorie additive $\mathcal{C}$, le morphisme canonique $\mathrm{Ext}^i_{\mathcal{F}_d(\mathcal{C})}(F,G) \to \mathrm{Ext}^i_{\mathcal{F}(\mathcal{C})}(F,G)$ est un isomorphisme seulement lorsque d est assez grand par rapport à i et au degré polynomial de F (ou de G). On renvoie à la remarque 3.12 pour une discussion plus développée à ce sujet.

La démonstration du théorème 1 est donnée à la section 3. Elle repose sur :

— une propriété d'annulation cohomologique très inspirée d'une propriété analogue dans les catégories de foncteurs sur une catégorie additive due au deuxième auteur (voir [11]) ;

— les propriétés de la filtration de l'anneau d'un produit direct de groupes libres par les puissances de son idéal d'augmentation, qui sont rappelées dans la section 2.

La dernière section est consacrée aux applications. Tout d'abord, on détermine la dimension homologique dans les catégories $\mathcal{F}_d(\mathbf{gr})$ de foncteurs fondamentaux : les puissances tensorielles de l'abélianisation (proposition 4.1) et les foncteurs de Passi (corollaire 4.3). Pour les premiers, on obtient :

Proposition 1. *Soient* $d \geq n > 0$ *des entiers. Le foncteur* $\mathfrak{a}^{\otimes n}$ *est de dimension homologique* $d-n$ *dans la catégorie* $\mathcal{F}_d(\mathbf{gr})$.

La finitude de ces dimensions constitue un phénomène rare dans les catégories de foncteurs polynomiaux et illustre la spécificité de $\mathcal{F}_d(\mathbf{gr})$.

On montre aussi le résultat suivant (proposition 4.6) :

Proposition 2. *Soient* $d > 0$ *un entier et* k *un sous-anneau de* $\mathbb{Q}$ *dans lequel* $d!$ *est inversible. La catégorie* $\mathcal{F}_d(\mathbf{gr};k)$ *est de dimension globale* $d-1$.

Dans la proposition 4.4, on considère le foncteur β_d adjoint à droite au foncteur $\mathcal{F}_d(\mathbf{gr}) \to \mathbb{Z}[\mathfrak{S}_d]\text{-}\mathbf{Mod}\quad F \mapsto cr_d(F)(\mathbb{Z},\dots,\mathbb{Z})$. On montre que ce foncteur, apparemment mystérieux (et en tout cas non explicite), mentionné dans [4], est exact, et induit donc des isomorphismes entre groupes d'extensions. Plus explicitement, nous obtenons :

Proposition 3. *Pour tout* $d \in \mathbb{N}$, *le foncteur* $\beta_d : \mathbb{Z}[\mathfrak{S}_d]\text{-}\mathbf{Mod} \to \mathcal{F}_d(\mathbf{gr})$ *est exact. Il induit des isomorphismes naturels*

$$\mathrm{Ext}^*_{\mathcal{F}(\mathbf{gr})}(\beta_d(M),\beta_d(N)) \simeq \mathrm{Ext}^*_{\mathcal{F}_d(\mathbf{gr})}(\beta_d(M),\beta_d(N)) \simeq \mathrm{Ext}^*_{\mathbb{Z}[\mathfrak{S}_d]}(M,N).$$

Là encore, il s'agit d'un phénomène notable dans les catégories de foncteurs polynomiaux, dont les liens avec les représentations des groupes symétriques sont bien compris pour ce qui concerne la classification des objets simples, par exemple, mais généralement mystérieux du point de vue cohomologique (notamment dans le cas d'une catégorie source additive, comme les groupes *abéliens* libres de rang fini).

Remerciements Les deux derniers auteurs souhaitent remercier les organisateurs du semestre intitulé *Grothendieck-Teichmuller Groups, Deformation and Operads* qui a eu lieu à l'Institut Isaac Newton de Cambridge en 2013 et pendant lequel ce travail à débuté ainsi que l'Institut pour son hospitalité et les excellentes conditions de travail qui y sont proposées. Le second auteur est reconnaissant envers l'université de Leicester pour lui avoir accordé un congé pour recherche.

1 Rappels sur les foncteurs polynomiaux

Soient k un anneau et $\mathcal{C}$ une petite catégorie. La catégorie $\mathcal{F}(\mathcal{C};k)$ est une catégorie abélienne qui hérite des propriétés de régularité de la catégorie but k-**Mod** ; en effet, les suites exactes, sommes, produits, etc. se testent au but. Si k est commutatif, on dispose d'une structure monoïdale symétrique notée $\otimes$ sur $\mathcal{F}(\mathcal{C};k)$ qui est le produit tensoriel sur k au but.
La catégorie $\mathcal{F}(\mathcal{C};k)$ possède suffisamment d'objets projectifs : en effet, le lemme de Yoneda montre que le foncteur

$$P_c^{\mathcal{C}} := k[\mathcal{C}(c,-)]$$

(où c est un objet de $\mathcal{C}$; $k[-]$ désigne le foncteur de linéarisation des ensembles vers k-**Mod**) représente l'évaluation en c, il est donc projectif (et de type fini), et les foncteurs $P_c^{\mathcal{C}}$ engendrent la catégorie $\mathcal{F}(\mathcal{C};k)$.
Si la catégorie $\mathcal{C}$ possède un objet nul 0 et des coproduits finis (notés ici $+$), on dispose dans $\mathcal{F}(\mathcal{C};k)$ d'une notion d'*effets croisés* et de *foncteurs polynomiaux* dans $\mathcal{F}(\mathcal{C};k)$ (ce cadre est exactement celui de [8], § 2.3) ; rappelons-en la définition. Si F est un foncteur de $\mathcal{F}(\mathcal{C};k)$, son n-ème effet croisé (*cross-effect* en anglais), où $n\in\mathbb{N}$, est le foncteur $cr_n(F)\in\mathcal{F}(\mathcal{C}^n;k)$ défini par

$$cr_n(F)(c_1,\dots,c_n)=\operatorname{Ker}\Big(F(c_1+\dots+c_n)\to\bigoplus_{i=1}^n F(c_1+\dots+\hat{c}_i+\dots+c_n)\Big)$$

(le chapeau signifie que le terme correspondant doit être omis ; les applications sont induites par les morphismes canoniques $c_1+\dots+c_n\to c_1+\dots+\hat{c}_i+\dots+c_n$ de $\mathcal{C}$ provenant de ce que cette catégorie possède des coproduits finis et un objet nul). Cet effet croisé définit un foncteur *exact* $cr_n:\mathcal{F}(\mathcal{C};k)\to\mathcal{F}(\mathcal{C}^n;k)$. L'exactitude peut se voir à partir des décompositions naturelles

$$F(c_1+\dots+c_n)\simeq\bigoplus_{1\le i_1<\dots<i_k\le n} cr_k(F)(c_{i_1},\dots,c_{i_k})$$

(cf. [8], proposition 2.4, par exemple). Une manière d'établir cette décomposition consiste à vérifier que $cr_n(F)(c_1, \dots, c_n)$ est le facteur direct de $F(c_1 + \dots + c_n)$ correspondant à l'idempotent $\sum_{I \subset \{1,\dots,n\}} (-1)^{\mathrm{Card}(I)} F(e_I)$, où e_I est l'idempotent de $c_1 + \dots + c_n$ donné par la projection sur les facteurs appartenant à I.

On dispose d'un isomorphisme canonique

$$cr_n(F)(c_{\sigma(1)}, \dots, c_{\sigma(n)}) \simeq cr_n(F)(c_1, \dots, c_n)$$

pour toute permutation $\sigma \in \mathfrak{S}_n$; en particulier, le groupe symétrique $\mathfrak{S}_n$ opère naturellement sur $cr_n(F)(c, \dots, c)$.

Comme $\mathcal{C}$ possède un objet nul, tout foncteur F de $\mathcal{F}(\mathcal{C}; k)$ se scinde de manière unique (à isomorphisme près) et naturelle en la somme directe d'un foncteur constant et d'un foncteur *réduit*, c'est-à-dire nul sur l'objet nul ; on note $\bar{F}$ ce foncteur réduit. Par ailleurs, comme $\mathcal{C}$ possède des coproduits finis, on dispose d'isomorphismes canoniques $P^{\mathcal{C}}_c \otimes P^{\mathcal{C}}_d \simeq P^{\mathcal{C}}_{c+d}$, lorsque k est commutatif. On vérifie classiquement que l'on dispose d'isomorphismes canoniques

$$\mathrm{Hom}_{\mathcal{F}(\mathcal{C};k)}(\bar{P}^{\mathcal{C}}_{c_1} \otimes \dots \otimes \bar{P}^{\mathcal{C}}_{c_n}, F) \simeq cr_n(F)(c_1, \dots, c_n).$$

Un foncteur F de $\mathcal{F}(\mathcal{C}; k)$ est dit *polynomial* de degré au plus n si $cr_{n+1}(F) = 0$. Cette condition implique l'annulation de $cr_i(F)$ pour tout entier $i > n$. On note $\mathcal{F}_n(\mathcal{C}; k)$ la sous-catégorie pleine de $\mathcal{F}(\mathcal{C}; k)$ constituée des foncteurs polynomiaux de degré au plus n. C'est une sous-catégorie épaisse de $\mathcal{F}(\mathcal{C}; k)$ stable par limites et colimites. Le produit tensoriel de deux foncteurs polynomiaux est polynomial (on suppose ici k commutatif), avec pour degré la somme des degrés des foncteurs initiaux. Le lecteur pourra trouver davantage de détails et de propriétés des foncteurs polynomiaux dans [8], § 2.

Venons-en au cas particulier de la catégorie source **gr**, qui possède un objet nul et des coproduits finis (donnés par le produit libre $*$). Pour alléger, on notera $P^{\mathbf{gr}}_n$ pour $P^{\mathbf{gr}}_{\mathbb{Z}^{*n}}$. On dispose d'isomorphismes canoniques $P^{\mathbf{gr}}_i \otimes P^{\mathbf{gr}}_j \simeq P^{\mathbf{gr}}_{i+j}$ et $P^{\mathbf{gr}}_n(G) \simeq k[G^n]$.

On notera simplement $\bar{P} = \bar{P}^{\mathbf{gr}}_1$ la partie réduite de $P^{\mathbf{gr}}_1$. Ainsi, $\bar{P}(G)$ n'est autre que l'idéal d'augmentation de la k-algèbre $k[G]$ du groupe libre G. Pour tous $d \in \mathbb{N}$ et $F \in \mathrm{Ob}\, \mathcal{F}(\mathbf{gr}; k)$, on dispose d'un isomorphisme naturel

$$\mathrm{Hom}_{\mathcal{F}(\mathbf{gr};k)}(\bar{P}^{\otimes n}, F) \simeq cr_n(F)(\mathbb{Z}, \dots, \mathbb{Z}).$$

2 Préliminaires sur les puissances de l'idéal d'augmentation d'un anneau de groupe

Nous donnons dans cette section plusieurs résultats concernant les puissances d'un idéal d'augmentation qui nous seront utiles dans la suite. La plupart de ces propriétés sont classiques. Une référence générale est [10].

Pour tout groupe G, on note $\mathcal{I}(G)$ l'idéal d'augmentation de l'anneau de groupe $\mathbb{Z}[G]$; pour tout $n \in \mathbb{N}$, on note $\mathcal{I}^n(G)$ la n-ème puissance de cet idéal. Pour tout

$n \in \mathbb{N}$, $\mathcal{I}^n$ définit un foncteur de la catégorie **Grp** des groupes vers la catégorie **Ab**. On dispose ainsi d'un anneau gradué $\mathrm{gr}(\mathbb{Z}[G]) := \bigoplus_{n\in\mathbb{N}} (\mathcal{I}^n/\mathcal{I}^{n+1})(G)$ naturel en G. En degré 1, on dispose d'un isomorphisme naturel $(\mathcal{I}/\mathcal{I}^2)(G) \simeq G_{ab}$ (abélianisation de G). En particulier, le produit induit des morphismes naturels $G_{ab}^{\otimes n} \to (\mathcal{I}^n/\mathcal{I}^{n+1})(G)$, qui sont toujours des épimorphismes (et sont des isomorphismes si G est libre — cf. remarque 2.2 infra).

Lemme 2.1. *Si G est un groupe libre, pour tous entiers naturels r et i, on dispose d'un isomorphisme naturel*

$$H_i(G; \mathcal{I}^r(G)) \simeq \begin{cases} (\mathcal{I}^r/\mathcal{I}^{r+1})(G) & si\ i = 0 \\ 0 & sinon. \end{cases}$$

Démonstration. On procède par récurrence sur r. L'assertion est triviale pour $r = 0$, on suppose donc $r > 0$ et le résultat établi pour $r - 1$. Dans la suite exacte de $\mathbb{Z}[G]$-modules à gauche

$$0 \to \mathcal{I}^r(G) \to \mathcal{I}^{r-1}(G) \to (\mathcal{I}^{r-1}/\mathcal{I}^r)(G) \to 0,$$

l'action de G est *triviale* sur $(\mathcal{I}^{r-1}/\mathcal{I}^r)(G)$, car la flèche de droite s'identifie à la projection de $\mathcal{I}^{r-1}(G)$ sur ses coïnvariants sous l'action de G — ce qui prouve déjà le résultat en degré homologique 0. Comme l'homologie de G est sans torsion sur $\mathbb{Z}$ et concentrée en degré 0 et 1, puisque G est libre, on en déduit que $H_i(G; (\mathcal{I}^{r-1}/\mathcal{I}^r)(G))$ est nul pour $i > 1$, isomorphe à $G_{ab} \otimes (\mathcal{I}^{r-1}/\mathcal{I}^r)(G)$ pour $i = 1$ et à $(\mathcal{I}^{r-1}/\mathcal{I}^r)(G)$ pour $i = 0$. La suite exacte longue d'homologie associée à la suite exacte courte précédente fournit donc le résultat d'annulation souhaité en degré homologique strictement positif. □

Remarque 2.2. Cette démonstration permet aussi de voir, en regardant la fin de ladite suite exacte longue, que le morphisme de liaison

$$G_{ab} \otimes (\mathcal{I}^{r-1}/\mathcal{I}^r)(G) \simeq H_1(G; (\mathcal{I}^{r-1}/\mathcal{I}^r)(G)) \to H_0(G; \mathcal{I}^r(G)) \simeq (\mathcal{I}^r/\mathcal{I}^{r+1})(G)$$

est un isomorphisme. Comme ce morphisme s'identifie au produit

$$(\mathcal{I}/\mathcal{I}^2)(G) \otimes (\mathcal{I}^{r-1}/\mathcal{I}^r)(G) \to (\mathcal{I}^r/\mathcal{I}^{r+1})(G),$$

cela fournit une démonstration de la propriété importante et classique (qui remonte à Magnus ; voir par exemple [10], chapitre VIII, *Theorem* 6.2) que l'anneau gradué $\mathrm{gr}(\mathbb{Z}[G])$ est naturellement isomorphe à l'algèbre tensorielle sur G_{ab} (pour G libre).

Lemme 2.3. *Soient G un groupe libre, r et i des entiers naturels. On dispose d'un isomorphisme naturel*

$$\mathrm{Tor}_i^{\mathbb{Z}[G]}(\mathcal{I}(G), \mathcal{I}^r(G)) \simeq \begin{cases} \mathcal{I}^{r+1}(G) & si\ i = 0 \\ 0 & sinon. \end{cases}$$

Démonstration. Utilisant la suite exacte longue d'homologie associée à la suite exacte courte de $\mathbb{Z}[G]$-modules à droite $0 \to \mathcal{I}(G) \to \mathbb{Z}[G] \to \mathbb{Z} \to 0$, on voit que $\mathrm{Tor}_i^{\mathbb{Z}[G]}(\mathcal{I}(G), \mathcal{I}^r(G))$ est isomorphe à $H_{i+1}(G; \mathcal{I}^r(G))$ lorsque $i > 0$, et l'on obtient une suite exacte

$$0 \to H_1(G; \mathcal{I}^r(G)) \to \mathrm{Tor}_0^{\mathbb{Z}[G]}(\mathcal{I}(G), \mathcal{I}^r(G)) \to \mathcal{I}^r(G) \to H_0(G; \mathcal{I}^r(G)) \to 0,$$

de sorte que le lemme 2.1 permet de conclure. □

Proposition 2.4. *Pour tout $r \in \mathbb{N}$, la résolution barre fournit un complexe de chaînes de foncteurs* **Grp** $\to$ **Ab** *du type*

$$\cdots \to \mathcal{I}^{\otimes(n+1)} \otimes \mathcal{I}^r \to \mathcal{I}^{\otimes n} \otimes \mathcal{I}^r \to \cdots \to \mathcal{I}^{\otimes 2} \otimes \mathcal{I}^r \to \mathcal{I} \otimes \mathcal{I}^r$$

dont la restriction aux groupes libres a une homologie isomorphe à $\mathcal{I}^{r+1}$ en degré 0 *et nulle en degré strictement positif.*

Démonstration. D'une manière générale, si A est un anneau augmenté, d'idéal d'augmentation $\bar{A}$, M un A-module à droite et N un A-module à gauche, on dispose d'un complexe de chaînes de groupes abéliens fonctoriel en A, M et N

$$\cdots \to M \otimes \bar{A}^{\otimes n} \otimes N \to M \otimes \bar{A}^{\otimes(n-1)} \otimes N \to \cdots \to M \otimes \bar{A} \otimes N \to M \otimes N$$

(complexe barre bilatère normalisé) dont l'homologie est naturellement isomorphe à $\mathrm{Tor}_*^A(M, N)$ si A et M sont sans torsion sur $\mathbb{Z}$.
La proposition s'obtient en prenant $A = \mathbb{Z}[G]$, $M = \mathcal{I}(G)$ et $N = \mathcal{I}^r(G)$ et en appliquant le lemme 2.3. □

Nous aurons besoin également d'examiner $\mathrm{gr}(\mathbb{Z}[G])$ lorsque G est un produit direct de groupes libres. Pour cela, on donne quelques propriétés générales simples sur l'effet du foncteur $\mathrm{gr}(\mathbb{Z}[-])$ sur un produit direct de groupes.
On va voir que ce foncteur est *exponentiel* (i.e. transforme les produits directs en produits tensoriels) sur les groupes G tels que $\mathrm{gr}(\mathbb{Z}[G])$ soit un groupe abélien sans torsion. (On a vu plus haut que les groupes libres possèdent cette propriété.) Remarquer que, si c'est le cas, alors les groupes abéliens $\mathcal{I}^n(G)$ et $\mathbb{Z}[G]/\mathcal{I}^n(G)$ sont également sans torsion.
Soient G et H deux groupes. Si M et N sont des sous-groupes des groupes abéliens $\mathbb{Z}[G]$ et $\mathbb{Z}[H]$ respectivement, nous noterons $M.N$ l'image de $M \otimes N$ dans $\mathbb{Z}[G] \otimes \mathbb{Z}[H] \simeq \mathbb{Z}[G \times H]$. Si M et N sont des idéaux bilatères de $\mathbb{Z}[G]$ et $\mathbb{Z}[H]$ respectivement, alors $M.N$ est un idéal bilatère de $\mathbb{Z}[G \times H]$.

Proposition 2.5. *Soient G et H deux groupes. Dans $\mathbb{Z}[G \times H]$, on a*

$$\mathcal{I}(G \times H) = \mathcal{I}(G).\mathbb{Z}[H] + \mathbb{Z}[G].\mathcal{I}(H)$$

et, plus généralement,

$$\mathcal{I}^r(G \times H) = \sum_{i+j=r} \mathcal{I}^i(G).\mathcal{I}^j(H)$$

pour tout $r \in \mathbb{N}$.

Démonstration. La première propriété est immédiate et la deuxième s'en déduit par récurrence sur r. □

Proposition 2.6. *Soient G et H deux groupes tels que les groupes abéliens* $\mathrm{gr}(\mathbb{Z}[G])$ *et* $\mathrm{gr}(\mathbb{Z}[H])$ *soient sans torsion.*

1. *Pour tout entier naturel n, on dispose d'un isomorphisme naturel*

$$(\mathcal{I}^n/\mathcal{I}^{n+1})(G\times H)\simeq\bigoplus_{i+j=n}(\mathcal{I}^i/\mathcal{I}^{i+1})(G)\otimes(\mathcal{I}^j/\mathcal{I}^{j+1})(H).$$

2. *Le groupe abélien* $\mathrm{gr}(\mathbb{Z}[G\times H])$ *est sans torsion.*
3. *Pour tous entiers $0\le t\le n$, on a*

$$\Big(\sum_{\substack{i+j=n\\ i<t}}\mathcal{I}^i(G).\mathcal{I}^j(H)\Big)\cap(\mathcal{I}^t(G).\mathcal{I}^{n-t}(H))=\mathcal{I}^t(G).\mathcal{I}^{n-t+1}(H)$$

dans $\mathbb{Z}[G\times H]$; ce groupe est naturellement isomorphe à $\mathcal{I}^t(G)\otimes\mathcal{I}^{n-t+1}(H)$.

Démonstration. La proposition 2.5 procure (sans aucune hypothèse sur G ni H) un épimorphisme naturel

$$\bigoplus_{i+j=n}(\mathcal{I}^i/\mathcal{I}^{i+1})(G)\otimes(\mathcal{I}^j/\mathcal{I}^{j+1})(H)\twoheadrightarrow(\mathcal{I}^n/\mathcal{I}^{n+1})(G\times H).$$

Par ailleurs, grâce à l'hypothèse faite sur G et H, les épimorphismes canoniques $\mathcal{I}^i(G)\otimes\mathcal{I}^j(H)\twoheadrightarrow\mathcal{I}^i(G).\mathcal{I}^j(H)$ sont des isomorphismes, et l'on a

$$\mathcal{I}^i(G).\mathcal{I}^{n-i}(H)\cap\mathcal{I}^j(G).\mathcal{I}^{n-j}(H)=\mathcal{I}^j(G).\mathcal{I}^{n-i}(H)$$

pour $i\le j$. On en déduit immédiatement la proposition. □

3 Démonstration du théorème 1

La classe $\mathcal{T}_n$ Soient $\mathcal{C}$ une petite catégorie possédant un objet nul et des coproduits finis. Pour $n\in\mathbb{N}\cup\{-1\}$, on note $\mathcal{T}_n(\mathcal{C})$ la classe des objets de $\mathcal{F}(\mathcal{C})$ possédant une résolution projective dont les termes sont des sommes directes de foncteurs du type $\bar{P}^{\mathcal{C}}_{c_1}\otimes\cdots\otimes\bar{P}^{\mathcal{C}}_{c_d}$, où $d>n$ est un entier et $c_1,\dots,c_d$ sont des objets de $\mathcal{C}$.
Le cas qui nous intéresse est celui des classes $\mathcal{T}_n(\mathbf{gr})$, qui seront simplement notées $\mathcal{T}_n$ par la suite. Toutefois, les propriétés formelles ci-dessous se montrent de la même façon, immédiate, dans le cas général ; elles sont laissées au lecteur. Noter que l'on a $\mathcal{T}_n(\mathcal{C})\supset\mathcal{T}_{n+1}(\mathcal{C})$ pour tout $n\in\mathbb{N}\cup\{-1\}$, que $\mathcal{T}_0(\mathcal{C})$ est constituée des foncteurs réduits et que $\mathcal{T}_{-1}(\mathcal{C})=\mathcal{F}(\mathcal{C})$.

Lemme 3.1. *Soient $n \in \mathbb{N} \cup \{-1\}$ et $0 \to F \to G \to H \to 0$ une suite exacte courte de $\mathcal{F}(\mathcal{C})$. Si deux des foncteurs F, G et H appartiennent à $\mathcal{T}_n(\mathcal{C})$, alors il en est de même pour le troisième.*
En particulier, si A et B sont deux sous-foncteurs d'un foncteur F de $\mathcal{F}(\mathcal{C})$, A et B appartenant à $\mathcal{T}_n(\mathcal{C})$, alors le sous-foncteur $A+B$ de F appartient à $\mathcal{T}_n(\mathcal{C})$ si et seulement s'il en est de même pour $A \cap B$.
Plus généralement, si on dispose d'une suite exacte longue

$$\cdots \to X_1 \to X_0 \to F \to 0$$

dans $\mathcal{F}(\mathcal{C})$ avec tous les X_i dans $\mathcal{T}_n(\mathcal{C})$, alors F appartient à $\mathcal{T}_n(\mathcal{C})$.

Proposition 3.2. *Si F et G appartiennent à $\mathcal{T}_i(\mathcal{C})$ et $\mathcal{T}_j(\mathcal{C})$ respectivement et que l'un de ces foncteurs prend des valeurs sans torsion sur $\mathbb{Z}$, alors $F \otimes G$ appartient à $\mathcal{T}_{i+j+1}(\mathcal{C})$.*

Corollaire 3.3. *Le produit tensoriel de $n+1$ foncteurs réduits de $\mathcal{F}(\mathcal{C})$ dont au moins n prennent des valeurs sans torsion sur $\mathbb{Z}$ appartient à $\mathcal{T}_n(\mathcal{C})$.*

Remarque 3.4. Ce corollaire est un analogue d'un résultat dû au deuxième auteur pour les catégories $\mathcal{F}(\mathcal{A})$, où $\mathcal{A}$ est additive, qui apparaît dans [11], et qui s'est avéré extrêmement utile en cohomologie des foncteurs.

Les classes $\mathcal{T}_n(\mathcal{C})$ nous serviront par l'intermédiaire du résultat suivant.

Proposition 3.5. *Soient $n \in \mathbb{N} \cup \{-1\}$, F un foncteur de $\mathcal{T}_n(\mathcal{C})$ à valeurs sans torsion sur $\mathbb{Z}$, k un anneau et A un foncteur de $\mathcal{F}_n(\mathcal{C};k)$. Alors $\mathrm{Ext}^*_{\mathcal{F}(\mathcal{C};k)}(F \otimes k, A) = 0$.*

Démonstration. Comme F est à valeurs sans torsion, $F \otimes k$ possède une résolution dont les termes sont des sommes directes de foncteurs du type $\bar{P}^{\mathcal{C}}_{c_1} \otimes \cdots \otimes \bar{P}^{\mathcal{C}}_{c_d} \otimes k$, avec $d > n$.
D'autre part, on dispose d'isomorphismes naturels

$$\begin{aligned}\mathrm{Hom}_{\mathcal{F}(\mathcal{C};k)}(\bar{P}^{\mathcal{C}}_{c_1} \otimes \cdots \otimes \bar{P}^{\mathcal{C}}_{c_d} \otimes k, A) \simeq \\ \simeq \mathrm{Hom}_{\mathcal{F}(\mathcal{C})}(\bar{P}^{\mathcal{C}}_{c_1} \otimes \cdots \otimes \bar{P}^{\mathcal{C}}_{c_d}, O_* A) \simeq cr_d(A)(c_1, \ldots, c_d)\end{aligned}$$

où $O : k\text{-}\mathbf{Mod} \to \mathbf{Ab}$ désigne le foncteur d'oubli, adjoint à droite à $- \otimes k$, de sorte que le foncteur de postcomposition $O_* : \mathcal{F}(\mathcal{C};k) \to \mathcal{F}(\mathcal{C})$ est adjoint à droite à la postcomposition par $- \otimes k$, et où le deuxième isomorphisme provient de la section 1. Le fait que les foncteurs exacts cr_d sont nuls sur $\mathcal{F}_n(\mathcal{C};k)$ pour $d > n$ permet de conclure. ☐

Les foncteurs K^d_n Pour tout $d \in \mathbb{N}$, le foncteur d'inclusion $\mathcal{F}_d(\mathbf{gr}) \to \mathcal{F}(\mathbf{gr})$ possède un adjoint à gauche q_d ; $q_d(F)$ est le plus grand quotient de F appartenant à $\mathcal{F}_d(\mathbf{gr})$. On renvoie à [8], § 2.3 pour plus de détails à ce sujet.

Notation 3.6. Étant donné deux entiers naturels n et d, on pose

$$K_n^d = \mathrm{Ker}(P_n^{\mathbf{gr}} \twoheadrightarrow q_d(P_n^{\mathbf{gr}}))$$

et l'on définit $K_n^{-1} = P_n^{\mathbf{gr}}$.

Proposition 3.7. *Pour tous $n \in \mathbb{N}$, $d \in \mathbb{N} \cup \{-1\}$ et tout objet G de $\mathbf{gr}$, les sous-groupes $K_n^d(G)$ et $\mathcal{I}^{d+1}(G^n)$ de $\mathbb{Z}[G^n]$ (auquel s'identifie $P_n^{\mathbf{gr}}(G)$) coïncident.*

Démonstration. Notons J_n^d le sous-foncteur de $P_n^{\mathbf{gr}}$ donné par $G \mapsto \mathcal{I}^{d+1}(G^n) \subset \mathbb{Z}[G^n] \simeq P_n^{\mathbf{gr}}(G)$. Alors le foncteur $P_n^{\mathbf{gr}}/J_n^d$ appartient à $\mathcal{F}_d(\mathbf{gr})$. Cela résulte de ce que, pour tout $r \in \mathbb{N}$, le foncteur $G \mapsto (\mathcal{I}^r/\mathcal{I}^{r+1})(G^n)$ (des groupes vers les groupes abéliens) est polynomial de degré (au plus) r, puisque c'est un quotient de $G \mapsto (G_{ab}^n)^{\otimes r}$ (voir le début de la section 2). On a donc $K_n^d \subset J_n^d$.
Montrons l'inclusion inverse. Pour tout foncteur F de $\mathcal{F}(\mathbf{gr})$ et tout objet G de $\mathbf{gr}$, on considère l'application naturelle

$$\kappa_d(F)(G) = \sum_{I \subset \{0,1,\dots,d\}} (-1)^{\mathrm{Card}(I)} F(p_I) : F(G^{*(d+1)}) \to F(G),$$

où $p_I \in \mathbf{gr}(G^{*(d+1)}, G) \simeq \mathrm{End}(G)^{d+1}$ est le morphisme dont la i-ème composante est l'identité si $i \in I$ et le morphisme trivial sinon. La transformation naturelle $\kappa_d(F)$ est nulle lorsque F appartient à $\mathcal{F}_d(\mathbf{gr})$, car elle se factorise par l'idempotent de $F(G^{*(d+1)})$ dont l'image est par définition l'effet croisé $cr_{d+1}(F)(G, \dots, G)$.
Par conséquent, comme $q_d(F)$ appartient à $\mathcal{F}_d(\mathbf{gr})$, la composée

$$F(G^{*(d+1)}) \xrightarrow{\kappa_d(F)(G)} F(G) \twoheadrightarrow q_d(F)(G)$$

est nulle, de sorte que le noyau de la projection $F \twoheadrightarrow q_d(F)$ contient l'image de $\kappa_d(F)$. En particulier, K_n^d contient l'image de $\kappa_d(P_n^{\mathbf{gr}})$. Si $(g_j^i)_{1 \le i \le n, 0 \le j \le d}$ est une famille d'éléments d'un groupe libre de rang fini G, on a

$$\kappa_d(P_n^{\mathbf{gr}})(G)\big([(g_0^i * \dots * g_d^i)_{1 \le i \le n}]\big) = \sum_{0 \le j_1 < \dots < j_r \le d} (-1)^r [(g_{j_1}^i \dots g_{j_r}^i)_{1 \le i \le n}]$$

qui est égal au produit

$$a_0 a_1 \dots a_d \quad \text{où} \quad a_j := [(g_j^i)_{1 \le i \le n}] - [1] \in \mathcal{I}(G^n) \subset \mathbb{Z}[G^n].$$

Comme $\mathcal{I}^{d+1}(G^n)$ est engendré par ces produits, cela montre que K_n^d contient J_n^d, d'où la conclusion. □

Remarque 3.8. Dans la démonstration précédente, le noyau de la projection $F \twoheadrightarrow q_d(F)$ est en fait *égal* à l'image de $\kappa_d(F)$ — cf. [8], définition 3.16 et proposition 3.17.

Lemme 3.9. *Pour tout $d \in \mathbb{N} \cup \{-1\}$, le foncteur K_1^d appartient à $\mathcal{T}_d$.*

Démonstration. On procède par récurrence sur d. Pour $d \leq 0$, c'est clair. Compte-tenu de la proposition 3.7, la proposition 2.4 peut se traduire par l'existence de suites exactes

$$\dots \to \bar{P}^{\otimes(n+1)} \otimes K_1^{d-1} \to \bar{P}^{\otimes n} \otimes K_1^{d-1} \to \dots$$
$$\dots \to \bar{P}^{\otimes 2} \otimes K_1^{d-1} \to \bar{P} \otimes K_1^{d-1} \to K_1^d \to 0$$

dans $\mathcal{F}(\mathbf{gr})$. Si K_1^{d-1} appartient à $\mathcal{T}_{d-1}$, alors $\bar{P}^{\otimes i} \otimes K_1^{d-1}$ appartient à $\mathcal{T}_d$ pour tout entier $i \geq 1$, donc K_1^d aussi (en utilisant le lemme 3.1), d'où le lemme. □

Proposition 3.10. *Pour tous $n \in \mathbb{N}$ et $d \in \mathbb{N} \cup \{-1\}$, le foncteur K_n^d appartient à $\mathcal{T}_d$.*

Démonstration. On montre par récurrence sur n que K_n^d appartient à $\mathcal{T}_d$ pour tout $d \in \mathbb{N} \cup \{-1\}$. Pour $n = 0$ il n'y a rien à faire ; compte-tenu du lemme précédent, on peut supposer $n > 1$ et l'assertion établie pour $n - 1$.
Dans la suite on conserve les notations de la fin de la section 2 et on fait sans cesse usage de l'identification canonique $P_n^{\mathbf{gr}}(G) = \mathbb{Z}[G^n]$ et de la proposition 3.7. On rappelle que, par le dernier point de la proposition 2.6, on a $K_t^i.K_s^j \simeq K_t^i \otimes K_s^j$.
Étant donné $d \in \mathbb{N} \cup \{-1\}$, on va montrer que, pour tout $t \in \{1, \dots, d+2\}$, le foncteur

$$\sum_{\substack{i+j=d+1 \\ 0 \leq i < t}} K_{n-1}^{i-1}.K_1^{j-1} \subset P_n^{\mathbf{gr}}$$

appartient à $\mathcal{T}_d$. Cela établira la proposition : en prenant $t = d+2$ et en utilisant la proposition 2.5, on en déduit bien que K_n^d appartient à $\mathcal{T}_d$.
Pour cela, on effectue une récurrence sur t. Lorsque $t = 1$, le foncteur considéré vaut $K_{n-1}^{-1}.K_1^d$. Par le lemme 3.9, K_1^d appartient à $\mathcal{T}_d$. On en déduit, par la proposition 3.2, que $K_{n-1}^{-1}.K_1^d \simeq K_{n-1}^{-1} \otimes K_1^d$ appartient à $\mathcal{T}_d$ (K_1^d étant un sous-foncteur de $P_1^{\mathbf{gr}}$, il est à valeurs $\mathbb{Z}$-plates). Supposons que pour $t - 1 \leq d + 1$ le foncteur considéré appartienne à $\mathcal{T}_d$. On a

$$\sum_{\substack{i+j=d+1 \\ 0 \leq i < t}} K_{n-1}^{i-1}.K_1^{j-1} = \Big(\sum_{\substack{i+j=d+1 \\ 0 \leq i < t-1}} K_{n-1}^{i-1}.K_1^{j-1}\Big) + K_{n-1}^{t-2}.K_1^{d-t+1}.$$

Par hypothèse de récurrence, le premier terme à droite de cette égalité appartient à $\mathcal{T}_d$. Le foncteur $K_{n-1}^{t-2}.K_1^{d-t+1} \simeq K_{n-1}^{t-2} \otimes K_1^{d-t+1}$ appartient également à $\mathcal{T}_d$ puisque K_1^{d-t+1} appartient à $\mathcal{T}_{d-t+1}$ par le lemme 3.9 et K_{n-1}^{t-2} appartient à $\mathcal{T}_{t-2}$ par hypothèse de récurrence et en appliquant la proposition 3.2. D'après le lemme 3.1, pour montrer que $\sum\limits_{\substack{i+j=d+1 \\ 0 \leq i < t}} K_{n-1}^{i-1}.K_1^{j-1}$ appartient à $\mathcal{T}_d$ il est donc équivalent de montrer que

$$(\sum_{\substack{i+j=d+1\\0\leq i<t-1}} K_{n-1}^{i-1}.K_1^{j-1}) \cap K_{n-1}^{t-2}.K_1^{d-t+1}$$

appartient à $\mathcal{T}_d$. Par la dernière partie de la proposition 2.6 on obtient que

$$(\sum_{\substack{i+j=d+1\\0\leq i<t-1}} K_{n-1}^{i-1}.K_1^{j-1}) \cap K_{n-1}^{t-2}.K_1^{d-t+1} = K_{n-1}^{t-2}.K_1^{d-t+2} \simeq K_{n-1}^{t-2} \otimes K_1^{d-t+2}.$$

Or K_1^{d-t+2} appartient à $\mathcal{T}_{d-t+2}$ d'après le lemme 3.9 et K_{n-1}^{t-2} appartient à $\mathcal{T}_{t-2}$ par hypothèse de récurrence. On déduit de la proposition 3.2 que $K_{n-1}^{t-2}.K_1^{d-t+2}$ appartient à $\mathcal{T}_{d+1} \subset \mathcal{T}_d$. Cela termine la démonstration. □

Démonstration du théorème 1. Pour tout $n \in \mathbb{N}$, on dispose d'isomorphismes d'adjonction

$$\mathrm{Hom}_{\mathcal{F}_d(\mathcal{C};k)}(q_d(P_n^{\mathbf{gr}}) \otimes k, F) \simeq \mathrm{Hom}_{\mathcal{F}_d(\mathcal{C})}(q_d(P_n^{\mathbf{gr}}), O_*F)$$

$$\simeq \mathrm{Hom}_{\mathcal{F}(\mathcal{C})}(P_n^{\mathbf{gr}}, O_*F) \simeq F(\mathbb{Z}^{*n})$$

où O_* désigne la postcomposition par le foncteur d'oubli $O : k\text{-}\mathbf{Mod} \to \mathbf{Ab}$, de sorte que les foncteurs $q_d(P_n^{\mathbf{gr}}) \otimes k$ forment un ensemble de générateurs projectifs de $\mathcal{F}_d(\mathbf{gr};k)$.
Par conséquent, il suffit de montrer le théorème 1 pour $F = q_d(P_n^{\mathbf{gr}}) \otimes k$. La suite exacte courte

$$0 \to K_n^d \to P_n^{\mathbf{gr}} \to q_d(P_n^{\mathbf{gr}}) \to 0$$

de $\mathcal{F}(\mathbf{gr})$ induit une suite exacte courte

$$0 \to K_n^d \otimes k \to P_n^{\mathbf{gr}} \otimes k \to q_d(P_n^{\mathbf{gr}}) \otimes k \to 0$$

dans $\mathcal{F}(\mathbf{gr};k)$ car $q_d(P_n^{\mathbf{gr}})$ prend des valeurs sans $\mathbb{Z}$-torsion. La suite exacte longue en Ext associée à cette suite exacte courte et le fait que $\mathrm{Ext}^*_{\mathcal{F}(\mathbf{gr};k)}(P_n^{\mathbf{gr}} \otimes k, G)$ et $\mathrm{Ext}^*_{\mathcal{F}_d(\mathbf{gr};k)}(q_d(P_n^{\mathbf{gr}}) \otimes k, G)$ sont nuls en degré cohomologique strictement positif montrent que la conclusion du théorème 1 équivaut à la nullité de $\mathrm{Ext}^*_{\mathcal{F}(\mathbf{gr};k)}(K_n^d \otimes k, G)$ lorsque G appartient à $\mathcal{F}_d(\mathbf{gr};k)$. Cela découle des propositions 3.10 et 3.5. □

Remarque 3.11. On peut employer les mêmes méthodes pour démontrer un résultat similaire au théorème 1 pour les groupes de torsion plutôt que les groupes d'extensions.

Remarque 3.12.

1. Si $\mathcal{C}$ est une petite catégorie additive, il est exceptionnel que l'inclusion de la sous-catégorie de foncteurs polynomiaux $\mathcal{F}_d(\mathcal{C})$ dans $\mathcal{F}(\mathcal{C})$ induise des isomorphismes entre *tous* les groupes d'extensions : dès le degré cohomologique $i = 3$, le morphisme canonique $\mathrm{Ext}^i_{\mathcal{F}_d(\mathcal{C})}(F, G) \to \mathrm{Ext}^i_{\mathcal{F}(\mathcal{C})}(F, G)$ peut cesser d'être un isomorphisme (même si F et G sont additifs et que l'on prend la colimite sur $d \in \mathbb{N}$). *Dans les bons cas* (avec une hypothèse de torsion bornée), ce morphisme est un isomorphisme lorsque d est assez grand par rapport à i et au degré de F ou de G. Ces résultats sont établis par le deuxième auteur dans [13] dans un cas particulier (F ou G additif et pas de torsion dans les groupes abéliens de morphismes de $\mathcal{C}$) et par le premier auteur dans [3] dans le cas général. Comme dans le présent travail, la démonstration de ces résultats nécessite d'analyser le gradué associé à la filtration de l'anneau d'un groupe (abélien, cette fois) par les puissances de l'idéal d'augmentation ; il faut toutefois aussi utiliser d'autres ingrédients, issus de la construction cubique de Mac Lane. Notons d'ailleurs que le cœur de notre démonstration consiste à montrer que la restriction des foncteurs $\mathcal{I}^{d+1}$ aux groupes libres appartient à $\mathcal{T}_d$; l'argument central de [13] (sa proposition 4.3) consiste à montrer une propriété analogue pour la restriction des foncteurs $\mathcal{I}^{d+1}$ aux groupes *abéliens* libres, propriété qui n'est valide qu'en degré cohomologique assez petit (et qui repose sur un argument d'idéal *quasi-régulier*).

2. Notons **mon** (un squelette de) la catégorie des *monoïdes* libres de rang fini. Le théorème 1 reste vrai si l'on remplace la catégorie source **gr** par **mon**. Une méthode pour le voir consiste à reprendre les arguments du présent article et les adapter à la catégorie $\mathcal{F}(\mathbf{mon})$. Une autre consiste à utiliser le résultat dû à Hartl et aux deuxième et troisième auteurs ([8], *Corollary* 5.38) selon lequel le foncteur de complétion en groupe $\alpha : \mathbf{mon} \to \mathbf{gr}$ induit pour chaque $d \in \mathbb{N}$ une équivalence de catégories $\mathcal{F}_d(\mathbf{gr}) \xrightarrow{\simeq} \mathcal{F}_d(\mathbf{mon})$. Dès lors, il suffit de voir que le foncteur α induit des isomorphismes entre groupes d'extensions
$$\mathrm{Ext}^*_{\mathcal{F}(\mathbf{gr})}(F, G) \xrightarrow{\simeq} \mathrm{Ext}^*_{\mathcal{F}(\mathbf{mon})}(F \circ \alpha, G \circ \alpha)$$
lorsque F et G sont polynomiaux. C'est en fait vrai si l'on suppose seulement que G est polynomial. Il suffit de le voir lorsque F est un foncteur projectif $P_n^{\mathbf{gr}}$; on peut alors utiliser un argument classique reposant sur le fait que le morphisme canonique d'un monoïde *libre* vers sa complétion en groupe induit un isomorphisme en homologie. Cet argument est donné en détail, dans un contexte abélien analogue, dans [1], théorème 3.3.

4 Applications

Notons $\mathfrak{a} : \mathbf{gr} \to \mathbf{Ab}$ le foncteur d'abélianisation. Ce foncteur joue un rôle fondamental dans la catégorie $\mathcal{F}(\mathbf{gr})$, où les calculs d'algèbre homologique sont grandement facilités par l'existence d'une résolution projective explicite de $\mathfrak{a}$. Celle-ci est donnée par la résolution barre (dont on tronque le degré nul : on utilise que l'homologie d'un groupe libre est naturellement isomorphe à son abélianisation en degré 1 et nulle en degré > 1), qui prend la forme :

$$\cdots \to P_{n+1}^{\mathbf{gr}} \to P_n^{\mathbf{gr}} \to \cdots \to P_2^{\mathbf{gr}} \to P_1^{\mathbf{gr}} \to \mathfrak{a} \to 0. \tag{1}$$

Cette résolution projective apparaît pour la première fois dans [9] (§ 5.A) ; elle est également utilisée de façon fondamentale dans [4].
En utilisant la résolution barre *normalisée*, on obtient une variante de la précédente résolution, également utile :

$$\cdots \to \bar{P}^{\otimes(n+1)} \to \bar{P}^{\otimes n} \to \cdots \to \bar{P}^{\otimes 2} \to \bar{P} \to \mathfrak{a} \to 0. \tag{2}$$

Proposition 4.1. *Soient $d \geq n > 0$ des entiers. Le foncteur $\mathfrak{a}^{\otimes n}$ est de dimension homologique $d - n$ dans la catégorie $\mathcal{F}_d(\mathbf{gr})$: on a $\mathrm{Ext}^i_{\mathcal{F}_d(\mathbf{gr})}(\mathfrak{a}^{\otimes n}, -) = 0$ pour $i > d - n$, tandis que le foncteur $\mathrm{Ext}^{d-n}_{\mathcal{F}_d(\mathbf{gr})}(\mathfrak{a}^{\otimes n}, -)$ n'est pas nul.*

Démonstration. En prenant le produit tensoriel de n copies de la résolution projective de $\mathfrak{a}$ dans $\mathcal{F}(\mathbf{gr})$ que procure la suite exacte (2) (dont tous les termes prennent des valeurs $\mathbb{Z}$-libres), on voit que $\mathfrak{a}^{\otimes n}$ possède une résolution projective qui, en degré i, est une somme directe de copies du foncteur $\bar{P}^{\otimes(i+n)}$. Comme $\mathrm{Hom}_{\mathcal{F}(\mathbf{gr})}(\bar{P}^{\otimes t}, F) \simeq cr_t(F)(\mathbb{Z}, \ldots, \mathbb{Z})$ est nul lorsque F est polynomial de degré $< t$ (cf. section 1), on en déduit $\mathrm{Ext}^i_{\mathcal{F}(\mathbf{gr})}(\mathfrak{a}^{\otimes n}, F) = 0$ pour F polynomial de degré $< n + i$.
Comme le foncteur $\mathrm{Ext}^i_{\mathcal{F}_d(\mathbf{gr})}(\mathfrak{a}^{\otimes n}, -)$ est la restriction à $\mathcal{F}_d(\mathbf{gr})$ du foncteur $\mathrm{Ext}^i_{\mathcal{F}(\mathbf{gr})}(\mathfrak{a}^{\otimes n}, -)$, d'après le théorème 1, on en déduit l'inégalité

$$\mathrm{hdim}_{\mathcal{F}_d(\mathbf{gr})}(\mathfrak{a}^{\otimes n}) \leq d - n.$$

L'inégalité inverse se déduit de la non-nullité de $\mathrm{Ext}^{d-n}_{\mathcal{F}(\mathbf{gr})}(\mathfrak{a}^{\otimes n}, \mathfrak{a}^{\otimes d})$, établie par le troisième auteur dans [15]. □

Remarque 4.2. L'analogue « abélien » de la proposition 4.1 n'est pas exact : notons $\mathbf{ab}$ la catégorie des groupes abéliens libres de rang fini. Dans la catégorie $\mathcal{F}_2(\mathbf{ab})$, le foncteur d'inclusion est de dimension homologique infinie et possède une résolution projective 4-périodique, comme on le déduit de la section 6 de l'article [13].
En revanche, on dispose quand même d'un résultat très similaire à la proposition 4.1 en remplaçant la catégorie source $\mathbf{gr}$ par une catégorie additive appropriée. Soient p un nombre premier et $\mathcal{F}(p)$ la catégorie des foncteurs des $\mathbb{F}_p$-espaces vectoriels de dimension finie vers les $\mathbb{F}_p$-espaces vectoriels. Le

foncteur d'inclusion est de dimension homologique finie $2p^{[\log_p(d)]}-2$ (où les crochets désignent la partie entière) dans la sous-catégorie $\mathcal{F}_d(p)$ des foncteurs polynomiaux de degré au plus d. Ce résultat est dû à Franjou et Smith ([7], § 4.2).

Les foncteurs $q_n(\bar{P})$ (appelés *foncteurs de Passi* dans [8]) jouent un rôle fondamental dans la catégorie $\mathcal{F}(\mathbf{gr})$ (la démonstration du théorème 1 en constitue une illustration) ; la proposition 4.1 permet facilement d'en calculer la dimension homologique.

Corollaire 4.3. *Soient $d \geq n > 0$ des entiers. Le foncteur $q_n(\bar{P})$ est de dimension homologique $d-n$ dans la catégorie $\mathcal{F}_d(\mathbf{gr})$.*

Démonstration. On dispose de suites exactes courtes

$$0 \to \mathfrak{a}^{\otimes n} \to q_n(\bar{P}) \to q_{n-1}(\bar{P}) \to 0$$

(cf. la proposition 3.7 et le début de la section 2), d'où des inégalités

$$\mathrm{hdim}_{\mathcal{F}_d(\mathbf{gr})}(q_{n-1}(\bar{P})) \leq \max\big(1+\mathrm{hdim}_{\mathcal{F}_d(\mathbf{gr})}(\mathfrak{a}^{\otimes n}), \mathrm{hdim}_{\mathcal{F}_d(\mathbf{gr})}(q_n(\bar{P}))\big).$$

On en déduit l'inégalité $\mathrm{hdim}_{\mathcal{F}_d(\mathbf{gr})}(q_n(\bar{P})) \leq d-n$ par récurrence descendante sur n, en utilisant la proposition 4.1 et le caractère projectif de $q_d(\bar{P})$ dans la catégorie $\mathcal{F}_d(\mathbf{gr})$.
L'inégalité inverse provient de ce que $\mathrm{Ext}^i_{\mathcal{F}(\mathbf{gr})}(q_n(\bar{P}), \mathfrak{a}^{\otimes(n+i)})$ est non nul (cf. [15]) et du théorème 1. □

Avant de donner la prochaine application du théorème 1, notons que l'action de $\mathfrak{S}_d$ sur $cr_d(F)(\mathbb{Z},\dots,\mathbb{Z})$ (voir la section 1) fournit, pour tout $d \in \mathbb{N}$, un foncteur $\mathrm{cr}_d : \mathcal{F}_d(\mathbf{gr}) \to \mathbb{Z}[\mathfrak{S}_d]$-$\mathbf{Mod}$ associant l'effet croisé $cr_d(F)(\mathbb{Z},\dots,\mathbb{Z})$ à F. Ce foncteur possède un adjoint à gauche α_d et un adjoint à droite β_d. Le foncteur α_d possède une expression explicite simple : $\alpha_d(M) = \mathfrak{a}^{\otimes d} \underset{\mathbb{Z}[\mathfrak{S}_d]}{\otimes} M$, mais il n'est pas exact (pour $d \geq 2$). En revanche, le foncteur β_d ne semble pas posséder d'expression simple. Néanmoins, on a :

Proposition 4.4. *Pour tout $d \in \mathbb{N}$, le foncteur $\beta_d : \mathbb{Z}[\mathfrak{S}_d]$-$\mathbf{Mod} \to \mathcal{F}_d(\mathbf{gr})$ est exact. Il induit des isomorphismes naturels*

$$\mathrm{Ext}^*_{\mathcal{F}(\mathbf{gr})}(\beta_d(M), \beta_d(N)) \simeq \mathrm{Ext}^*_{\mathcal{F}_d(\mathbf{gr})}(\beta_d(M), \beta_d(N)) \simeq \mathrm{Ext}^*_{\mathbb{Z}[\mathfrak{S}_d]}(M, N).$$

Démonstration. Pour voir que β_d est exact, il suffit de vérifier que son adjoint à gauche cr_d envoie les générateurs projectifs $q_d(P_n^{\mathbf{gr}})$ de $\mathcal{F}_d(\mathbf{gr})$ (où n parcourt $\mathbb{N}$) sur des $\mathbb{Z}[\mathfrak{S}_d]$-modules projectifs. Comme cr_d est exact et tue les foncteurs de degré strictement inférieur à d, il prend la même valeur sur $q_d(P_n^{\mathbf{gr}})$ et le noyau Q_n^d de la projection $q_d(P_n^{\mathbf{gr}}) \twoheadrightarrow q_{d-1}(P_n^{\mathbf{gr}})$. Or on a $Q_1^d \simeq \mathfrak{a}^{\otimes d}$ (car le gradué de l'anneau d'un groupe libre est isomorphe à l'algèbre tensorielle de son abélianisation — cf. section 2 ; on utilise également la proposition 3.7).

En général, en utilisant le premier point de la proposition 2.6 (et encore la proposition 3.7), on obtient :

$$Q_n^d \simeq \bigoplus_{i_1+\dots+i_n=d} \mathfrak{a}^{\otimes i_1} \otimes \dots \otimes \mathfrak{a}^{\otimes i_n} \simeq \bigoplus_{i_1+\dots+i_n=d} \mathfrak{a}^{\otimes d}.$$

Comme cr_d envoie le foncteur $\mathfrak{a}^{\otimes d}$ sur le $\mathbb{Z}[\mathfrak{S}_d]$-module $\mathbb{Z}[\mathfrak{S}_d]$, cela démontre l'exactitude de β_d.
L'adjonction entre les foncteurs exacts β_d et cr_d se propage aux groupes d'extensions :

$$\mathrm{Ext}^*_{\mathcal{F}_d(\mathbf{gr})}(F, \beta_d(N)) \simeq \mathrm{Ext}^*_{\mathbb{Z}[\mathfrak{S}_d]}(\mathrm{cr}_d(F), N).$$

Comme la coünité $\mathrm{cr}_d \beta_d \to \mathrm{Id}$ est un isomorphisme (cf. [4], § 2, par exemple), en utilisant le théorème 1, on en déduit la dernière assertion de l'énoncé. □

Remarque 4.5.

1. En reprenant la démonstration précédente, il est facile de donner une expression explicite de $\beta_d(M)(G)$ fonctorielle en M, *mais pas en* $G \in \mathrm{Ob}\,\mathbf{gr}$.
2. La proposition 4.4 contraste encore avec la situation de $\mathcal{F}(\mathbf{ab})$. On y dispose de même d'un foncteur

$$\mathcal{F}_d(\mathbf{ab}) \to \mathbb{Z}[\mathfrak{S}_d]\text{-}\mathbf{Mod} \quad F \mapsto cr_d(F)(\mathbb{Z}, \dots, \mathbb{Z})$$

qui possède des adjoints de chaque côté. L'adjoint à gauche est très similaire au foncteur α_d évoqué avant la proposition (il est donné par $M \mapsto T^d \underset{\mathbb{Z}[\mathfrak{S}_d]}{\otimes} M \simeq (T^d \otimes M)_{\mathfrak{S}_d}$, où T^d désigne la d-ème puissance tensorielle). L'adjoint à droite est analogue à l'adjoint à gauche (il est donné par $M \mapsto (T^d \otimes M)^{\mathfrak{S}_d}$) ; c'est donc encore un foncteur explicite mais *non exact* si $d > 1$. La référence originelle pour cette question classique est [12].

La proposition 4.4 montre que les catégories $\mathcal{F}_d(\mathbf{gr})$ (resp. $\mathcal{F}_d(\mathbf{gr}; \mathbb{F}_p)$, où p est un nombre premier) sont de dimension globale infinie dès que $d \geq 2$ (resp. $d \geq p$), comme l'algèbre de groupe $\mathbb{Z}[\mathfrak{S}_d]$ (resp. $\mathbb{F}_p[\mathfrak{S}_d]$).
Notre dernier résultat montre qu'en revanche, les catégories $\mathcal{F}_d(\mathbf{gr}; \mathbb{Q})$ (ou plus généralement $\mathcal{F}_d(\mathbf{gr}; k)$, où k est un sous-anneau de $\mathbb{Q}$ où assez d'entiers sont inversés) sont de dimension globale finie.

Proposition 4.6. *Soient $d > 0$ un entier et k un sous-anneau de $\mathbb{Q}$ dans lequel $d!$ est inversible. La catégorie $\mathcal{F}_d(\mathbf{gr}; k)$ est de dimension globale $d - 1$: on a $\mathrm{Ext}^i_{\mathcal{F}_d(\mathbf{gr};k)} = 0$ pour $i \geq d$, tandis que le foncteur $\mathrm{Ext}^{d-1}_{\mathcal{F}_d(\mathbf{gr};k)}$ n'est pas nul.*

Démonstration. On montre d'abord l'inégalité $\mathrm{gldim}\,\mathcal{F}_d(\mathbf{gr}; k) \leq d - 1$. Par le théorème 1, cela équivaut à dire que la restriction à $\mathcal{F}_d(\mathbf{gr}; k)^{op} \times \mathcal{F}_d(\mathbf{gr}; k)$ du foncteur $\mathrm{Ext}^i_{\mathcal{F}(\mathbf{gr};k)}$ est nulle pour $i \geq d$. Pour cela, on va montrer par récurrence sur $d \in \mathbb{N}^*$ l'assertion suivante :

$$\forall i \geq n \geq d \quad \forall F \in \mathrm{Ob}\,\mathcal{F}_d(\mathbf{gr};k) \quad \forall G \in \mathrm{Ob}\,\mathcal{F}_n(\mathbf{gr};k) \qquad \mathrm{Ext}^i(F,G) = 0$$

(on omet dans la suite l'indice $\mathcal{F}(\mathbf{gr};k)$ pour les groupes d'extensions). On suppose donc le résultat établi pour les entiers strictement inférieurs à d.
Les résultats de structure de la catégorie $\mathcal{F}_d(\mathbf{gr};k)$ donnés dans [4] (voir la section 2 et la démonstration de la proposition 5.3 de cet article) impliquent que, pour tout foncteur F de $\mathcal{F}_d(\mathbf{gr};k)$, il existe une représentation M du groupe symétrique $\mathfrak{S}_d$ et un morphisme $u : F \to \mathfrak{a}^{\otimes d} \underset{k[\mathfrak{S}_d]}{\otimes} M$ dont le noyau N et le conoyau C appartiennent à $\mathcal{F}_{d-1}(\mathbf{gr};k)$. Pour $d = 1$, on peut même supposer que u est un isomorphisme, puisque la partie constante des foncteurs se scinde. L'hypothèse de récurrence montre que $\mathrm{Ext}^i(N,G)$ et $\mathrm{Ext}^i(C,G)$ sont nuls pour G dans $\mathcal{F}_n(\mathbf{gr};k)$ et $i \geq n \geq d-1$. Par ailleurs, comme l'anneau $k[\mathfrak{S}_d]$ est semi-simple à cause de l'hypothèse faite sur k, le foncteur $T := \mathfrak{a}^{\otimes d} \underset{k[\mathfrak{S}_d]}{\otimes} M$ est somme directe de facteurs directs de $\mathfrak{a}^{\otimes d}$. La proposition 4.1 montre donc que $\mathrm{Ext}^i(T,G) = 0$ pour G dans $\mathcal{F}_n(\mathbf{gr};k)$ et $i > n-d \geq 0$. On en déduit $\mathrm{Ext}^i(F,G) = 0$ pour G dans $\mathcal{F}_n(\mathbf{gr};k)$ et $i \geq n \geq d$, ce qui termine la démonstration de l'inégalité $\mathrm{gldim}\,\mathcal{F}_d(\mathbf{gr};k) \leq d-1$.
L'inégalité $\mathrm{gldim}\,\mathcal{F}_d(\mathbf{gr};k) \geq d-1$ se déduit de ce que le groupe abélien $\mathrm{Ext}^{d-1}_{\mathcal{F}(\mathbf{gr})}(\mathfrak{a}, \mathfrak{a}^{\otimes d})$ est non seulement non nul, mais aussi sans torsion (cf. [15]). □

Remarque 4.7. Contrairement à ce qui advient pour la plupart des autres résultats du présent article, dont les analogues sur **ab** sont plus difficiles à montrer que sur **gr** (voir notamment la remarque 4.2), quand ils ne sont pas faux, la situation est plus simple pour $\mathcal{F}_d(\mathbf{ab};\mathbb{Q})$, qui est une catégorie *semi-simple* pour tout $d \in \mathbb{N}$ (et ce résultat classique est aisé à prouver).

Références

[1] Aurélien Djament. Sur l'homologie des groupes unitaires à coefficients polynomiaux. *J. K-Theory*, 10(1) :87–139, 2012.

[2] Aurélien Djament. Décomposition de Hodge pour l'homologie stable des groupes d'automorphismes des groupes libres https ://hal.archives-ouvertes.fr/hal-01214646, 2015.

[3] Aurélien Djament. Groupes d'extensions et foncteurs polynomiaux. *J. Lond. Math. Soc. (2)*, 92(1) :63–88, 2015.

[4] Aurélien Djament and Christine Vespa. Sur l'homologie des groupes d'automorphismes des groupes libres à coefficients polynomiaux. *Comment. Math. Helv.*, 90(1) :33–58, 2015.

[5] Samuel Eilenberg and Saunders Mac Lane. On the groups $H(\Pi, n)$. II. Methods of computation. *Ann. of Math. (2)*, 60 :49–139, 1954.

[6] Vincent Franjou, Eric M. Friedlander, Teimuraz Pirashvili, and Lionel Schwartz. *Rational representations, the Steenrod algebra and functor homology*, volume 16 of *Panoramas et Synthèses [Panoramas and Syntheses]*. Société Mathématique de France, Paris, 2003.

[7] Vincent Franjou and Jeffrey H. Smith. A duality for polynomial functors. *J. Pure Appl. Algebra*, 104(1) :33–39, 1995.

[8] Manfred Hartl, Teimuraz Pirashvili, and Christine Vespa. Polynomial functors from algebras over a set-operad and nonlinear Mackey functors. *Int. Math. Res. Not. IMRN*, (6) :1461–1554, 2015.

[9] Mamuka Jibladze and Teimuraz Pirashvili. Cohomology of algebraic theories. *J. Algebra*, 137(2) :253–296, 1991.

[10] Inder Bir S. Passi. *Group rings and their augmentation ideals*, volume 715 of *Lecture Notes in Mathematics*. Springer, Berlin, 1979.

[11] T. I. Pirashvili. Higher additivizations. *Trudy Tbiliss. Mat. Inst. Razmadze Akad. Nauk Gruzin. SSR*, 91 :44–54, 1988.

[12] T. I. Pirashvili. Polynomial functors. *Trudy Tbiliss. Mat. Inst. Razmadze Akad. Nauk Gruzin. SSR*, 91 :55–66, 1988.

[13] Teimuraz Pirashvili. Polynomial approximation of Ext and Tor groups in functor categories. *Comm. Algebra*, 21(5) :1705–1719, 1993.

[14] Teimuraz Pirashvili. Dold-Kan type theorem for Γ-groups. *Math. Ann.*, 318(2) :277–298, 2000.

[15] Christine Vespa. Extensions between functors from groups. arXiv :1511.03098.

Aurélien Djament
CNRS
Laboratoire de Mathématiques
Jean Leray (UMR 6629)
aurelien.djament@univ-nantes.fr

Teimuraz Pirashvili
Department of Mathematics
University of Leicester
University Road
Leicester LE1 7RH
UK
tp59@leicester.ac.uk

Christine Vespa
Institut de Recherche
Mathématique Avancée
Université de Strasbourg
vespa@math.unistra.fr

On F-Crystalline Representations

Bryden Cais, Tong Liu[1]

Received: May 5, 2015
Revised: November 26, 2015

Communicated by Don Blasius

Abstract. We extend the theory of Kisin modules and crystalline representations to allow more general coefficient fields and lifts of Frobenius. In particular, for a finite and totally ramified extension $F/\mathbb{Q}_p$, and an arbitrary finite extension K/F, we construct a general class of infinite and totally wildly ramified extensions K_∞/K so that the functor $V \mapsto V|_{G_{K_\infty}}$ is fully-faithfull on the category of F-crystalline representations V. We also establish a new classification of F-Barsotti-Tate groups via Kisin modules of height 1 which allows more general lifts of Frobenius.

2010 Mathematics Subject Classification: Primary 14F30,14L05
Keywords and Phrases: F-crystalline representations, Kisin modules

1. Introduction

Let k be a perfect field of characteristic p with ring of Witt vectors $W := W(k)$, write $K_0 := W[1/p]$ and let K/K_0 be a finite and totally ramified extension. We fix an algebraic closure $\overline{K}$ of K and set $G_K := \mathrm{Gal}(\overline{K}/K)$. The theory of *Kisin modules* and its variants, pioneered by Breuil in [Bre98] and later developed by Kisin [Kis06], provides a powerful set of tools for understanding Galois-stable $\mathbb{Z}_p$-lattices in $\mathbb{Q}_p$-valued semistable G_K-representations, and has been a key ingredient in many recent advances (*e.g.* [Kis08], [Kis09a], [Kis09b]). Throughout this theory, the non-Galois "Kummer" extension K_∞/K—obtained by adjoining to K a compatible system of choices $\{\pi_n\}_{n\geq 1}$ of p^n-th roots of a uniformizer π_0 in K—plays central role. The theory of Kisin modules closely parallels Berger's work [Ber02], in which the cyclotomic extension of K replaces K_∞, and can be thought of as a "K_∞-analogue" of the theory of Wach modules developed by Wach [Wac96], Colmez [Col99] and Berger [Ber04]. Along these

[1]The second author is partially supported by a Sloan fellowship and NSF grant DMS-1406926.

lines, Kisin and Ren [KR09] generalized the theory of Wach modules to allow the cyclotomic extension of K to be replaced by an arbitrary Lubin–Tate extension.

This paper grew out of a desire to better understand the role of K_∞ in the theories of Breuil and Kisin and related work, and is an attempt to realize Kisin modules and the modules of Wach and Kisin–Ren as "specializations" of a more general theory. To describe our main results, we first fix some notation.

Let $F \subseteq K$ be a subfield which is finite over $\mathbb{Q}_p$ with residue field k_F of cardinality $q = p^s$. Choose a power series

$$f(u) := a_1 u + a_2 u^2 + \cdots \in \mathcal{O}_F[\![u]\!]$$

with $f(u) \equiv u^q \bmod \mathfrak{m}_F$ and a uniformizer π of K with monic minimal polynomial $E := E(u)$ over $F_0 := K_0 \cdot F$. We set $\pi_0 := \pi$ and we choose $\underline{\pi} := \{\pi_n\}_{n\geq 1}$ with $\pi_n \in \overline{K}$ satisfying $f(\pi_n) = \pi_{n-1}$ for $n \geq 1$. The extension $K_{\underline{\pi}} := \bigcup_{n\geq 0} K(\pi_n)$ (called a *Frobenius iterate extension* in [CD15]) is an infinite and totally wildly ramified extension of K which need not be Galois. We set $G_{\underline{\pi}} := \mathrm{Gal}(\overline{K}/K_{\underline{\pi}})$.

Define $\mathfrak{S} := W[\![u]\!]$ and put $\mathfrak{S}_F = \mathcal{O}_F \otimes_{W(k_F)} \mathfrak{S} = \mathcal{O}_{F_0}[\![u]\!]$. We equip $\mathfrak{S}_F$ with the (unique continuous) Frobenius endomorphism φ which acts on $W(k)$ by the canonical q-power Witt-vector Frobenius, acts as the identity on $\mathcal{O}_F$, and sends u to $f(u)$. A *Kisin module of E-height r* is a finite free $\mathfrak{S}_F$-module $\mathfrak{M}$ endowed with φ-semilinear endomorphism $\varphi_{\mathfrak{M}} : \mathfrak{M} \to \mathfrak{M}$ whose linearization $1 \otimes \varphi : \varphi^*\mathfrak{M} \to \mathfrak{M}$ has cokernel killed by E^r.

When $F = \mathbb{Q}_p$ and $f(u) = u^p$ (which we refer to as the *classical situation* in the following), Kisin's theory [Kis06] attaches to any G_{K_∞}-stable $\mathbb{Z}_p$-lattice T in a semistable G_K-representation V with Hodge–Tate weights in $\{0, \ldots, r\}$ a unique Kisin module $\mathfrak{M}$ of height r satisfying $T \simeq T_{\mathfrak{S}}(\mathfrak{M})$ (see §3.3 for the definition of $T_{\mathfrak{S}}$). Using this association, Kisin proves that the restriction functor $V \to V|_{G_{K_\infty}}$ is fully faithful when restricted to the category of crystalline representations, and that the category of Barsotti–Tate groups over $\mathcal{O}_K$ is *anti-equivalent* to the category of Kisin modules of height 1.

In this paper, we extend much of the framework of [Kis06] to allow general F and $f(u)$, though for simplicity we will restrict ourselves to the case that $q = p$, or equivalently that $F/\mathbb{Q}_p$ is totally ramified, and that $f(u)$ is a monic degree-p polynomial. When we extend our coefficients from $\mathbb{Q}_p$ to F, we must further restrict ourselves to studying *F-crystalline representations*, which are defined following ([KR09]): Let V be a finite dimensional F-vector space with continuous F-linear action of G_K. If V is crystalline (when viewed as a $\mathbb{Q}_p$-representation) then $D_{\mathrm{dR}}(V)$ is naturally an $F \otimes_{\mathbb{Q}_p} K$-module and one has a decomposition $D_{\mathrm{dR}}(V) = \prod_{\mathfrak{m}} D_{\mathrm{dR}}(V)_{\mathfrak{m}}$, with $\mathfrak{m}$ running over the maximal ideals of $F \otimes_{\mathbb{Q}_p} K$. We say that V is *F-crystalline* if the induced filtration on $D_{\mathrm{dR}}(V)_{\mathfrak{m}}$ is trivial unless $\mathfrak{m}$ corresponds to the canonical inclusion $F \subset K$.

Theorem 1.0.1. *Let V be an F-crystalline representation with Hodge-Tate weights in $\{0,\ldots,r\}$ and $T \subset V$ a $G_{\underline{\pi}}$-stable $\mathcal{O}_F$-lattice. Then there exists a Kisin module $\mathfrak{M}$ of E-height r satisfying $T_{\mathfrak{S}}(\mathfrak{M}) \simeq T$.*

Writing v_F for the normalized valuation of $\overline{K}$ with $v_F(F) = \mathbb{Z}$, apart from the classical situation $f(u) = u^p$ of Kisin, the above theorem is also known when $v_F(a_1) = 1$, which corresponds to the Lubin–Tate cases covered by the work of [KR09]. An important point of our formalism is that $\mathfrak{M}$ may in general *not* be unique for a fixed lattice T: our general construction produces as special cases the φ-modules over $\mathfrak{S}_F$ which occur in the theory of Wach modules and its generalizations [KR09], so without the additional action of a Lubin–Tate group Γ, one indeed *does not expect* these Kisin modules to be uniquely determined; (*cf.* Example 3.3.7). This is of course quite different from the classical situation. Nonetheless, we prove the following version of Kisin's "full-faithfulness" result. Writing $\mathrm{Rep}_F^{F\text{-cris},r}(G)$ for the category of F-crystalline representations with Hodge-Tate weights in $\{0,\ldots,r\}$ and $\mathrm{Rep}_F(G_{\underline{\pi}})$ for the category of F-linear representations of $G_{\underline{\pi}}$, we prove:

Theorem 1.0.2. *Assume that $\varphi^n(f(u)/u)$ is not a power of E for all $n \geq 0$ and that $v_F(a_1) > r$, where $f(u) = a_1u + a_2u^2 + \cdots$. Then the restriction functor $\mathrm{Rep}_F^{F\text{-cris},r}(G) \rightsquigarrow \mathrm{Rep}_F(G_{\underline{\pi}})$ induced by $V \mapsto V|_{G_{\underline{\pi}}}$ is fully faithfull.*

Although Beilinson and Tavares Ribeiro [BTR13] have given an almost elementary proof of Theorem 1.0.2 in the classical situation $F = \mathbb{Q}_p$ and $f(u) = u^p$, their argument relies crucially on an explicit description of the Galois closure of K_∞/K. For more general F and f, we have no idea what the Galois closure of $K_{\underline{\pi}}/K$ looks like, and describing it in any explicit way seems to be rather difficult in general.

It is natural to ask when two different choices f and f' of p-power Frobenius lifts and corresponding compatible sequences $\underline{\pi} = \{\pi_n\}_n$ and $\underline{\pi}' = \{\pi'_n\}$ in $\overline{K}$ yield the *same* subfield $K_{\underline{\pi}} = K_{\underline{\pi}'}$ of $\overline{K}$. We prove that this is *rare* in the following precise sense: if $K_{\underline{\pi}} = K_{\underline{\pi}'}$, then the lowest degree terms of f and f' coincide, up to multiplication by a unit in $\mathcal{O}_F$; see Proposition 3.1.3. It follows that there are *infinitely many* distinct $K_{\underline{\pi}}$ for which Theorem 1.0.2 applies. We also remark that any Frobenius–iterate extension $K_{\underline{\pi}}$ as above is an infinite and totally wildly ramified *strictly APF* extension in the sense of Wintenberger [Win83]. We therefore think of Theorem 1.0.2 as confirmation of the philosophy that "crystalline p-adic representations are the p-adic analogue of unramified ℓ-adic representations[2]," since Theorem 1.0.2 is obvious if "crystalline" is replaced with "unramified" throughout (or equivalently in the special case $r = 0$). More generally, given F and $r \geq 0$, it is natural to ask for a characterization of *all* infinite and totally wildly ramified strictly APF extensions L/K for which

[2]This philosophy is perhaps best evinced by the p-adic analogue of the good reduction theorem of Néron–Ogg–Shafarevich, which asserts that an abelian variety A over a p-adic field K has good reduction if and only if its p-adic Tate module V_pA is a crystalline representation of G_K [CI99, Theorem 1].

restriction of F-crystalline representations of G_K with Hodge–Tate weights in $\{0, \ldots, r\}$ to G_L is fully–faithful. We believe that there should be a deep and rather general phenomenon which deserves further study.

While the condition that $v_F(a_1) > r$ is really essential in Theorem 1.0.2 (see Example 4.5.9), we suspect the conclusion is still valid if we remove the assumption that $\varphi^n(f(u)/u)$ is not a power of E for all $n \geq 0$. However, we have only successfully removed this assumption when $r = 1$, thus generalizing Kisin's classification of Barsotti–Tate groups:

Theorem 1.0.3. *Assume $v_F(a_1) > 1$. Then the category of Kisin modules of height 1 is equivalent to the category of F-Barsotti-Tate groups over $\mathcal{O}_K$.*

Here, an *F-Barsotti-Tate group* is a Bartotti–Tate group H over $\mathcal{O}_K$ with the property that the p-adic Tate module $V_p(H) = \mathbb{Q}_p \otimes_{\mathbb{Z}_p} T_p(H)$ is an F-crystalline representation. We note that when $F = \mathbb{Q}_p$, Theorem 1.0.3 is proved (by different methods) in [CL14].

Besides providing a natural generalization of Kisin's work and its variants as well as a deeper understanding of some of the finer properties of crystalline p-adic Galois representations, we expect that our theory will have applications to the study of potentially Barsotti–Tate representations. More precisely, suppose that T is a finite free $\mathcal{O}_F$-linear representation of G_K with the property that $T|_{G_{K'}}$ is Barsotti-Tate for some finite extension K'/K. If K'/K is not tamely ramified then it is well-known that it is in general difficult to construct "descent data" for the Kisin module $\mathfrak{M}$ associated to $T|_{G_{K'}}$ in order to study T (see the involved computations in [BCDT01]). However, suppose that we can select $f(u)$ and π_0 such that $K' \subseteq K(\pi_n)$ for some n. Then, as in the theory of Kisin–Ren [KR09] (see also [BB10]), we expect the appropriate descent data on $\mathfrak{M}$ to be *much easier* to construct in this "adapted" situation, and we hope this idea can be used to study the reduction of T.

Now let us sketch the ideas involved in proving the above theorems and outline the organization of this paper. For any $\mathbb{Z}_p$-algebra A, we set $A_F := A \otimes_{\mathbb{Z}_p} \mathcal{O}_F$. In order to connect $\mathfrak{S}_F$ to Galois representations, we must first embed $\mathfrak{S}_F$ as a Frobenius-stable subring of $W(R)_F$, which we do in §2.1 following [CD15]. In the following subsection, we collect some useful properties of this embedding and study some "big rings" inside $B^+_{\mathrm{cris},F}$. Contrary to the classical situation, the Galois closure of $K_{\underline{\pi}}$ appears in general to be rather mysterious. Nonetheless, in §2.3 we are able to establish some basic results on the G_K-conjugates of $u \in \mathfrak{S}_F \subseteq W(R)_F$ which are just barely sufficient for the development of our theory. Following Fontaine [Fon90], and making use of the main result of [CD15], in §3 we establish a classification of $G_{\underline{\pi}}$-representations via étale φ-modules and Kisin modules. In the end of §3, we apply these considerations to prove that the functor $T_{\mathfrak{S}}$ is fully faithful under the assumption that $\varphi^n(f(u)/u)$ is not a power of E for any n.

The technical heart of this paper is §4. In §4.1, we define F-crystalline representations and attach to each F-crystalline representation V a filtered φ-module

$D_{\mathrm{cris},F}(V)$ (we warn the reader that the filtration of $D_{\mathrm{cris},F}(V)$ is slightly different from that of $D_{\mathrm{cris}}(V)$). Following [Kis06], in §4.2 we then associate to $D = D_{\mathrm{cris}}(V)$ a φ-module $\mathcal{M}(D)$ over $\mathfrak{O}$ (here we use $\mathfrak{O}$ for the analogue of $\mathcal{O}$—the ring of rigid-analytic functions on the open unit disk—in Kisin's work). A shortcoming in our situation is that we do not in general know how to define a reasonable differential operator N_∇, even at the level of the ring $\mathfrak{O}$. Consequently, our $\mathcal{M}(D)$ only has a Frobenius structure, in contrast to the classical (and Lubin–Tate) situation in which $\mathcal{M}(D)$ is also equipped with a natural N_∇-structure. Without such an N_∇-structure, there is no way to follow Kisin's (or Berger's) original strategy to prove that the scalar extension of $\mathcal{M}(D)$ to the Robba ring is pure of slope zero, which is key to showing that there exists a Kisin module $\mathfrak{M}$ such that $\mathfrak{O} \otimes_{\mathfrak{S}_F} \mathfrak{M} \simeq \mathcal{M}(D)$. We bypass this difficulty by appealing to the fact that $\mathcal{M}(D)$ is known to be pure of slope zero in the classical situation of Kisin as follows: letting a superscript of "c" denote the data in the classical situation and using the fact that both $\mathcal{M}(D)$ and $\mathcal{M}^c(D)$ come from the same D, we prove that $\widetilde{B}_\alpha \otimes_{\mathfrak{O}} \mathcal{M}(D) \simeq \widetilde{B}_\alpha \otimes_{\mathfrak{O}^c} \mathcal{M}^c(D)$ as φ-modules for a certain period ring $\widetilde{B}_\alpha$ that contains the ring $\widetilde{B}^+_{\mathrm{rig},F}$. It turns out that this isomorphism can be descended to $\widetilde{B}^+_{\mathrm{rig},F}$. Since Kedlaya's theory of the slope filtration is unaffected by base change from the Robba ring to $\widetilde{B}^+_{\mathrm{rig},F}$, it follows that $\mathcal{M}(D)$ is of pure slope 0 as this is the case for $\mathcal{M}^c(D)$ thanks to [Kis06]. With this crucial fact in established, we are then able to prove Theorem 1.0.1 along the same lines as [Kis06]. If our modules came equipped with a natural N_∇-structure, the full faithfulness of the functor $V \mapsto V|_{G_{\underline{\pi}}}$ would follow easily from the full faithfulness of $T_{\mathfrak{S}}$. But without such a structure, we must instead rely heavily on the existence of a *unique* φ-equivariant section $\xi : D(\mathfrak{M}) \to \mathfrak{O}_\alpha \otimes \varphi^*\mathfrak{M}$ to the projection $\varphi^*\mathfrak{M} \twoheadrightarrow \varphi^*\mathfrak{M}/u\varphi^*\mathfrak{M}$, where $D(\mathfrak{M}) = (\varphi^*\mathfrak{M}/u\varphi^*\mathfrak{M})[1/p]$. The hypothesis $v_F(a_1) > r$ of Theorem 1.0.2 guarantees the existence and uniqueness of such a section ξ. With these preparations, we finally prove Theorem 1.0.2 in §4.5.

In §5, we establish Theorem 1.0.3: the equivalence between the category of Kisin modules of height 1 and the category of F-Barsotti-Tate groups over $\mathcal{O}_K$. Here we adapt the ideas of [Liu13b] to prove that the functor $\mathfrak{M} \mapsto T_{\mathfrak{S}}(\mathfrak{M})$ is an equivalence between the category of Kisin module of height 1 and the category of G_K-stable $\mathcal{O}_F$-lattices in F-crystalline representations with Hodge-Tate weights in $\{0, 1\}$. The key difficulty is to extend the $G_{\underline{\pi}}$-action on $T_{\mathfrak{S}}(\mathfrak{M})$ to a G_K-action which gives $T_{\mathfrak{S}}(\mathfrak{M})[1/p]$ the structure of an F-crystalline representation. In the classical situation, this is done using the (unique) monodromy operator N on $S \otimes_{\mathfrak{S}} \varphi^*\mathfrak{M}$ (see §2.2 in [Liu13b]). Here again, we are able to sidestep the existence of a monodromy operator to construct a (unique) G_K-action on $W(R)_F \otimes_{\mathfrak{S}_F} \mathfrak{M}$ which is compatible with the additional structures (see Lemma 5.1.1), and this is enough for us to extend the given $G_{\underline{\pi}}$-action to a G_K-action on $T_{\mathfrak{S}}(\mathfrak{M})$. As this paper establishes analogues of many of the results of [Kis06] in our more general context, it is natural ask to what extent the entire theory of [Kis06] can be developed in this setting. To that end, we list several interesting (some quite promising) questions for this program in the last section.

Acknowledgements: It is pleasure to thank Laurent Berger, Kiran Kedlaya and Ruochuan Liu for very helpful conversations and correspondence.

Notation. Throughout this paper, we reserve φ for the Frobenius operator, adding appropriate subscripts as needed for clarity: for example, $\varphi_{\mathfrak{M}}$ denotes the Frobenius map on $\mathfrak{M}$. We will always drop these subscripts when there is no danger of confusion. Let S be a ring endowed with a Frobenius lift φ_S and M an S-module. We always write $\varphi^*M := S \otimes_{\varphi_S, S} M$. Note that if $\varphi_M : M \to M$ is a φ_S-semilinear endomorphism, then $1 \otimes \varphi_M : \varphi^* M \to M$ is an S-linear map. We reserve $f(u) = u^p + a_{p-1}u + \cdots + a_1 u \equiv u^p \bmod \mathfrak{m}_F$ for the polynomial over $\mathcal{O}_F$ giving our Frobenius lift $\varphi(u) := f(u)$ as in the introduction. For any discretely valued subfield $E \subseteq \overline{K}$, we write v_E for the normalized p-adic valuation of $\overline{K}$ with $v_E(E) = \mathbb{Z}$, and for convenience will simply write $v := v_{\mathbb{Q}_p}$. If A is a $\mathbb{Z}_p$-module, we set $A_F := A \otimes_{\mathbb{Z}_p} \mathcal{O}_F$ and $A[1/p] := A \otimes_{\mathbb{Z}_p} \mathbb{Q}_p$. For simplicity, we put $G = G_K := \mathrm{Gal}(\overline{K}/K)$ and $G_{\underline{\pi}} := \mathrm{Gal}(\overline{K}/K_{\underline{\pi}})$. We write $\mathrm{M}_d(S)$ for the ring of $d \times d$-matrices over S and I_d for the $d \times d$-identity matrix.

2. Period rings

In this section, we introduce and study the various "period rings" which will play a central role in the development of our theory.

As in the introduction, we fix a perfect field k of characteristic p with ring of Witt vectors $W := W(k)$, as well as a finite and totally ramified extension K of $K_0 := W[1/p]$. Let F be a subfield of K, which is finite and totally ramified over $\mathbb{Q}_p$, and put $F_0 := K_0 F \subset K$. Choose uniformizers π of $\mathcal{O}_K$ and ϖ of $\mathcal{O}_F$, and let $E := E(u) \in \mathcal{O}_{F_0}[u]$ be the monic minimal polynomial of π over F_0. We set $e := [K : K_0]$, and put $e_0 := [K : F_0]$ and $e_F := [F : \mathbb{Q}_p]$. Fix a polynomial $f(u) = u^p + a_{p-1}u^{p-1} + \cdots + a_1 u \in \mathcal{O}_F[u]$ satisfying $f(u) \equiv u^p \bmod \varpi$, and recursively choose $\pi_n \in \overline{K}$ with $f(\pi_n) = \pi_{n-1}$ for $n \geq 1$ where $\pi_0 := \pi$. Set $K_{\underline{\pi}} := \bigcup_{n \geq 0} K(\pi_n)$ and $G_{\underline{\pi}} := \mathrm{Gal}(\overline{K}/K_{\underline{\pi}})$, and recall that for convenience we write $G = G_K := \mathrm{Gal}(\overline{K}/K)$.

Recall that $\mathfrak{S} = W[\![u]\!]$, and that we equip the scalar extension $\mathfrak{S}_F$ with the semilinear Frobenius endomorphism $\varphi : \mathfrak{S}_F \to \mathfrak{S}_F$ which acts on W as the unique lift of the p-power Frobenius map on k, acts trivially on $\mathcal{O}_F$, and sends u to $f(u)$. The first step in our classification of F-crystalline G_K-representations by Kisin modules over $\mathfrak{S}_F$ is to realize this ring as a Frobenius stable subring of $W(R)_F$, which we do in the following subsection.

2.1. $\mathfrak{S}_F$ as a subring of $W(R)_F$. As usual, we put $R := \varprojlim_{x \mapsto x^p} \mathcal{O}_{\overline{K}}/(p)$, equipped with its natural coordinate-wise action of G. It is well-known that the natural reduction map

$$\varprojlim_{x\to x^p} \mathcal{O}_{\overline{K}}/(p) \to \varprojlim_{x\to x^p} \mathcal{O}_{\overline{K}}/(\varpi)$$

is an isomorphism, so $\{\pi_n \bmod \varpi\}_{n\geq 0}$ defines an element $\underline{\pi} \in R$. Furthermore, writing $\mathbb{C}_K$ for the completion of $\overline{K}$, reduction modulo p yields a multiplicative bijection $\varprojlim_{x\to x^p} \mathcal{O}_{\mathbb{C}_K} \simeq R$, and for any $x \in R$ we write $(x^{(n)})_{n\geq 0}$ for the p-power compatible sequence in $\varprojlim_{x\to x^p} \mathcal{O}_{\mathbb{C}_K}$ corresponding to x under this identification. We write $[x] \in W(R)$ for the Techmüller lift of $x \in R$, and denote by $\theta : W(R) \to \mathcal{O}_{\mathbb{C}_K}$ the unique lift of the projection $R \twoheadrightarrow \mathcal{O}_{\mathbb{C}_K}/(p)$ which sends $\sum_n p^n[x_n]$ to $\sum_n p^n x^{(0)}$. By definition, B^+_{dR} is the $\mathrm{Ker}(\theta)$-adic completion of $W(R)[1/p]$, so θ naturally extends to B^+_{dR}. For any subring $B \subset B^+_{\mathrm{dR}}$, we define $\mathrm{Fil}^i B := (\mathrm{Ker}\theta)^i \cap B$.

There is a canonical section $\overline{K} \hookrightarrow B^+_{\mathrm{dR}}$, so we may view F as a subring of B^+_{dR}, and in this way we obtain embeddings $W(R)_F \hookrightarrow B^+_{\mathrm{cris},F} \hookrightarrow B^+_{\mathrm{dR}}$. Define $\theta_F := \theta|_{W(R)_F}$. One checks that $W(R)_F$ is ϖ-adically complete and that every element of $W(R)_F$ has the form $\sum_{n\geq 0}[a_n]\varpi^n$ with $a_n \in R$. The map θ_F carries $\sum_{n\geq 0}[a_n]\varpi^n$ to $\sum_{n\geq 0} a_n^{(0)}\varpi^n \in \mathcal{O}_{\mathbb{C}_K}$ (see Def. 3.8 and Prop. 3.9 of [CD15]).

Lemma 2.1.1. *There is a unique set-theoretic section $\{\cdot\}_f : R \to W(R)_F$ to the reduction modulo ϖ map which satisfies $\varphi(\{x\}_f) = f(\{x\}_f)$ for all $x \in R$.*

Proof. This is[3] [Col02, Lemme 9.3]. Using the fact that $f(u) \equiv u^p \bmod \varpi$, one checks that the endomorphism $f \circ \varphi^{-1}$ of $W(R)_F$ is a ϖ-adic contraction, so that for any lift $\widetilde{x} \in W(R)$ of $x \in R$, the limit

$$\{x\}_f := \lim_{n\to\infty} (f\circ\varphi^{-1})^{(n)}(\widetilde{x})$$

exists in $W(R)_F$ and is the unique fixed point of $f \circ \varphi^{-1}$, which uniquely characterizes it independent of our choice of $\widetilde{x}$. □

From Lemma 2.1.1 we obtain a unique continuous embedding $\iota : \mathfrak{S}_F \hookrightarrow W(R)_F$ of $\mathcal{O}_F$-algebras with $\iota(u) := \{\underline{\pi}\}_f$. Via ι, we henceforth identify $\mathfrak{S}_F$ with a φ-stable $\mathcal{O}_F$-subalgebra of $W(R)_F$ on which we have $\varphi(u) = f(u)$.

Example 2.1.2 (Cyclotomic case). Let $\{\zeta_{p^n}\}_{n\geq 0}$ be a compatible system of primitive p^n-th roots of unity. Let $K = \mathbb{Q}_p(\zeta_p)$ with $\pi = \zeta_p - 1$ and take $F = \mathbb{Q}_p$ with $f(u) = (u+1)^p - 1$. Choosing $\pi_n = \zeta_{p^{n+1}} - 1$, we obtain $K_{\underline{\pi}} := \mathbb{Q}_p(\mu_{p^\infty})$. It is obvious that $\underline{\epsilon}_1 := (\zeta_{p^n})_{n\geq 1} \in R$. In this case, $\iota(u) = [\underline{\epsilon}_1] - 1 \in W(R)$.

Recall that R has the structure of a valuation ring via $v_R(x) := v(x^{(0)})$, where v is the normalized p-adic valuation of $\mathbb{C}_K$ with $v(\mathbb{Z}_p) = \mathbb{Z}$.

Lemma 2.1.3. *We have $\theta_F(u) = \pi$ and E generates* $\mathrm{Ker}(\theta_F) = \mathrm{Fil}^1 W(R)_F$.

[3] In the version of Colmez's article available from his website, it is Lemme 8.3.

Proof. The first assertion is [Col02, Lemme 9.3]. To compute $\theta_F(\{\underline{\pi}\}_f)$, we first choose $[\underline{\pi}]$ as our lift of $\underline{\pi}$ to $W(R)$, and compute

$$\theta_F(\{\underline{\pi}\}_f) = \theta_F\left(\lim_{n\to\infty} f^{(n)}\varphi^{(-n)}([\underline{\pi}])\right) = \lim_{n\to\infty} f^{(n)}\theta_F([\underline{\pi}^{p^{-n}}]) = \lim_{n\to\infty} f^{(n)}(\underline{\pi}^{(n)})$$

But $\underline{\pi}^{(n)} \equiv \pi_n \bmod \varpi$, so

$$f^{(n)}(\underline{\pi}^{(n)}) \equiv f^{(n)}(\pi_n) \equiv \pi \bmod \varpi^{n+1},$$

which gives the claim. Now certainly $\theta_F(E(u)) = E(\pi) = 0$, so $E = E(u)$ lies in $\mathrm{Fil}^1 W(R)_F$. Since $E \equiv \underline{\pi}^{e_0} \bmod \varpi$, we conclude that

$$v_R(E \bmod \varpi) = e_0 v_R(\underline{\pi}) = e_0 v(\pi) = v(\varpi),$$

whence E generates $\mathrm{Ker}(\theta_F) = \mathrm{Fil}^1 W(R)_F$ thanks to [Col02, Prop. 8.3]. $\square$

Now let us recall the construction of $B^+_{\max}$ and $\widetilde{B}^+_{\mathrm{rig}}$ from Berger's paper [Ber02]. Let ξ be a generator of $\mathrm{Fil}^1 W(R)$. By definition,

$$B^+_{\max} := \left\{ \sum_{n\geq 0} a_n \frac{\xi^n}{p^n} \in B^+_{\mathrm{dR}} \;\middle|\; a_n \in W(R)[1/p] \text{ and } \lim_{n\to\infty} a_n = 0 \right\}.$$

and $\widetilde{B}^+_{\mathrm{rig}} := \bigcap_{n\geq 1} \varphi^n(B^+_{\max})$.

Write $\mathfrak{u} := [\underline{\pi}]$. The discussion before Proposition 8.14 in [Col02] shows:

Lemma 2.1.4.

$$\begin{aligned} B^+_{\max,F} &= \left\{ \sum_{n\geq 0} a_n \frac{E^n}{\varpi^n} \in B^+_{\mathrm{dR}} \;\middle|\; a_n \in W(R)_F[1/p] \text{ and } \lim_{n\to\infty} a_n = 0 \right\} \\ &= \left\{ \sum_{n\geq 0} a_n \frac{\mathfrak{u}^{e_0 n}}{\varpi^n} \in B^+_{\mathrm{dR}} \;\middle|\; a_n \in W(R)_F[1/p] \text{ and } \lim_{n\to\infty} a_n = 0 \right\}. \end{aligned}$$

We can now prove the following result, which will be important in §4.4:

Lemma 2.1.5. *Let $x \in B^+_{\max,F}$, and suppose that $xE^r = \varphi^m(y)$ holds for some $y \in B^+_{\max,F}$. Then $x = \varphi^m(y')$ with $y' \in B^+_{\max,F}$.*

Proof. By Lemma 2.1.4, we may write $y = \sum_n b_n \frac{\mathfrak{u}^{e_0 n}}{\varpi^n}$ with $b_n \in W(R)_F[1/p]$ converging to 0. Write $E = E(u) = \mathfrak{u}^{e_0} + \varpi z$ with $z \in W(R)_F$. We then have

$$\varphi^m(y) = \sum_{n=0}^{\infty} \varphi^m(b_n) \frac{\mathfrak{u}^{e_0 p^m n}}{\varpi^n} = \sum_{n=0}^{\infty} \varphi^m(b_n) \frac{(E - \varpi z)^{p^m n}}{\varpi^n} = \sum_{n=0}^{\infty} c_n \frac{E^{p^m n}}{\varpi^n}$$

with $c_n \in W(R)_F[1/p]$ converging to 0. By Lemma 2.1.3, E is a generator of $\mathrm{Fil}^1 W(R)_F$, so defininig $s := 1 + \max\{n \mid p^m n < r\}$, it follows that $\sum_{n=0}^{s-1} c_n \frac{E^{p^m n}}{\varpi^n}$

is divisible by E^r in $W(R)_F[1/p]$ so we may write $\sum_{n=0}^{s-1} c_n \frac{E^{p^m n}}{\varpi^n} = E^r x_0$ for some $x_0 \in W(R)[1/p]$. Without loss of generality, replacing x by $x - x_0$, gives

$$x = \sum_{n=s}^{\infty} c_n \frac{E^{p^m n - r}}{\varpi^n} = \sum_{n=s}^{\infty} d_{n-s} \frac{E^{p^m(n-s)}}{\varpi^{n-s}} = \sum_{n=0}^{\infty} d_n \frac{E^{p^m n}}{\varpi^n}$$

with $d_{n-s} = c_n \frac{E^{p^m s - r}}{\varpi^s}$. Using again the equality $E = \mathfrak{u}^{e_0} + \varpi z$, we then obtain $x = \sum_{n=0}^{\infty} e_n \frac{\mathfrak{u}^{e_0 p^m n}}{\varpi^n}$ with $e_n \in W(R)_F[1/p]$ converging to 0. We now have $x = \varphi^m(y')$ for $y' := \sum_{n=0}^{\infty} f_n \frac{\mathfrak{u}^{e_0 n}}{\varpi^n}$ with $f_n = \varphi^{-m}(e_n)$. As $f_n \in W(R)_F[1/p]$ converges to 0, we conclude that $y' \in B^+_{\max,F}$, as desired. $\square$

2.2. Some subrings of $B^+_{\mathrm{cris},F}$. For a subinterval $I \subset [0,1)$, we write $\mathfrak{O}_I$ for the subring of $F_0(\!(u)\!)$ consisting of those Laurent series which converge for those $x \in \mathbb{C}_K$ with $|x| \in I$, and we will simply write $\mathfrak{O} = \mathfrak{O}_{[0,1)}$. Let $\widetilde{B}_\alpha := W(R)_F[\![\frac{E^p}{\varpi}]\!][1/p] \subset B^+_{\mathrm{cris},F}$. We claim that $\mathrm{Fil}^n \widetilde{B}_\alpha = E^n \widetilde{B}_\alpha$. To see this, set $c = \lceil \frac{n}{p} \rceil$ and $n = pc - s$ with $0 \le s < p$. For any $x \in \mathrm{Fil}^n \widetilde{B}_\alpha$, we write $x = \sum_{i=0}^{\infty} a_i \frac{E^{pi}}{\varpi^i}$ with $a_i \in W(R)_F[1/p]$ converging to 0 in $W(R)_F[1/p]$. Since $x \in \mathrm{Fil}^n B^+_{\mathrm{dR}}$, $\sum_{i=0}^{c-1} a_i \frac{E^{pi}}{\varpi^i} = E^n x_0$ with $x_0 \in W(R)_F[1/p]$. It suffices to show that $x - x_0 = E^n y$ with $y \in \widetilde{B}_\alpha$. Now

$$y = \sum_{i \ge c}^{\infty} a_i \frac{E^{pi-n}}{\varpi^i} = \sum_{i \ge c}^{\infty} a_i E^s \varpi^{-c} \left(\frac{E^{p(i-c)}}{\varpi^{i-c}} \right) \in B^+_{\mathrm{dR}}.$$

As a_i converges to 0 in $W(R)_F[1/p]$, so does $a_i E^s \varpi^{-c}$, whence y lies in $\widetilde{B}_\alpha$.

Lemma 2.2.1. *There are canonical inclusions of rings $\mathfrak{O} \subset \widetilde{B}^+_{\mathrm{rig},F} \subset \widetilde{B}_\alpha$.*

Proof. We first show that $\mathfrak{O} \subset \widetilde{B}^+_{\mathrm{rig},F}$. For any $h(u) = \sum_{n=0}^{\infty} a_n u^n \in \mathfrak{O}$, we have to show that $h_m(u) = \sum_{n=0}^{\infty} \varphi^{-m}(a_n u^n)$ is in $B^+_{\max,F}$ for all $m \ge 0$. Writing $u = \mathfrak{u} + \varpi z$ with $\mathfrak{u} = [\underline{\pi}]$ and $z \in W(R)_F$, we have $\varphi^{-m}(u) = \mathfrak{u}^{p^{-m}} + \varpi z^{(m)}$ with $z^{(m)} = \varphi^{-m}(z) \in W(R)_F$. Setting $a_n^{(m)} := \varphi^{-m}(a_n) \in F_0$, we then have

$$h_m(u) = h(\mathfrak{u}^{\frac{1}{p^m}} + \varpi z^{(m)}) = \sum_{k=0}^{\infty} \frac{h^{(k)}(\mathfrak{u}^{\frac{1}{p^m}})}{k!} (\varpi z^{(m)})^k,$$

for $h^{(k)}$ the k-th derivative of $\widetilde{h}(X) := \sum\limits_{n=0}^{\infty} a_n^{(m)} X^n$. Therefore,

$$h_m(u) = \sum_{n=0}^{\infty}\left(\sum_{k=0}^{\infty}\binom{k+n}{k} a_{n+k}^{(m)}(\varpi z^{(m)})^k\right)\mathfrak{u}^{\frac{n}{p^m}}.$$

Since $h(u) \in \mathfrak{O}_{[0,1)}$, we have $\lim\limits_{n\to\infty} |a_n^{(m)}| r^n = 0$ for any $r < 1$. It follows that the inner sum $\sum\limits_{k=0}^{\infty}\binom{k+n}{k} a_{n+k}^{(m)}(\varpi z^{(m)})^k$ converges to $b_n \in W(R)_F[1/p]$. Since $\lim\limits_{n\to\infty} |a_n^{(m)}| r^n = 0$ for $r = |\varpi^{\frac{1}{e_0 p^m}}| \geq |\varpi|$, for any $\epsilon > 0$, there exists N so that $|a_{n+k}^{(m)} \varpi^{\frac{n}{e_0 p^m}} \varpi^k| < \epsilon$ for any $n > N$ and $k \geq 0$. This implies that $b_n \varpi^{\frac{n}{e_0 p^m}}$ converges to 0 in $W(R)_F$. We may therefore write

$$h_m(u) = \sum_{n=0}^{\infty} b_n \mathfrak{u}^{\frac{n}{p^m}} = \sum_{n=0}^{\infty} b_n \varpi^{\frac{n}{e_0 p^m}} \frac{(\mathfrak{u}^{e_0})^{\frac{n}{e_0 p^m}}}{\varpi^{\frac{n}{e_0 p^m}}},$$

and Lemma 2.1.4 implies that $h_m(u) \in B^+_{\max,F}$, so $\mathfrak{O} \subset \widetilde{B}^+_{\mathrm{rig},F}$ as desired.

To show that $\widetilde{B}^+_{\mathrm{rig},F} \subset \widetilde{B}_\alpha$, we first observe that

$$\widetilde{B}_\alpha = W(R)_F[\![\frac{u^{e_0 p}}{\varpi}]\!][1/p] = W(R)_F[\![\frac{\mathfrak{u}^{e_0 p}}{\varpi}]\!][1/p]. \tag{2.2.1}$$

For any $x \in \widetilde{B}^+_{\mathrm{rig},F}$, we may write $x = \varphi(y)$ with $y = \sum\limits_{n=0}^{\infty} a_n \frac{\mathfrak{u}^{e_0 n}}{\varpi^n} \in B^+_{\max,F}$, and we see that $x = \sum\limits_{n=0}^{\infty} \varphi(a_n) \frac{\mathfrak{u}^{e_0 p n}}{\varpi^n}$ indeed lies in $\widetilde{B}_\alpha$ by (2.2.1). □

Finally let us record the following technical lemma: recall that our Frobenius lift on $\mathfrak{S}_F$ is determined by $\varphi(u) := f(u)$, with $f(u) = u^p + a_{p-1}u^{p-1} + \cdots + a_1 u$. We define $\mathfrak{O}_\alpha := \mathfrak{S}_F[\![\frac{u^{e_0 p}}{\varpi}]\!][1/p] \subset \widetilde{B}_\alpha$.

Lemma 2.2.2. *Suppose that $\varpi^{r+1} | a_1$ in $\mathcal{O}_F$. Then there exists $h_i^{(n)}(u) \in \mathcal{O}_F[u]$ such that*

$$f^{(n)}(u) = \sum_{i=0}^{n} h_{n-i}^{(n)}(u) u^{2^{n-i}} \varpi^{(r+1)i}.$$

In particular, $\varphi^n(u)/\varpi^{rn}$ converges to 0 in $\mathfrak{O}_\alpha$.

Proof. We proceed by induction on $m = n$. When $m = 1$, we may write

$$f(u) = u^p + a_{p-1}u^{p_1} + \cdots + a_1 u = u^2 h(u) + b_0 \varpi^{r+1} u \quad \text{with} \quad b_0 \in \mathcal{O}_F. \tag{2.2.2}$$

Supposing that the assertion holds for $m = n$ and using (2.2.2) we compute

$$
\begin{aligned}
f^{(n+1)}(u) &= \sum_{i=0}^{n} h_{n-i}^{(n)}(f(u)) f(u)^{2^{n-i}} \varpi^{(r+1)i} \\
&= \sum_{i=0}^{n} h_{n-i}^{(n)}(f(u))(u^2 h(u) + b_0 \varpi^{r+1} u)^{2^{n-i}} \varpi^{(r+1)i} \\
&= \sum_{i=0}^{n} h_{n-i}^{(n)}(f(u)) \left(\sum_{k=0}^{2^{n-i}} \binom{2^{n-i}}{k} (u^2 h(u))^{2^{n-i}-k} (b_0 \varpi^{r+1} u)^k \right) \varpi^{(r+1)i} \\
&= \sum_{i=0}^{n} \sum_{k=0}^{2^{n-i}} \left(h_{n-i}^{(n)}(f(u)) \binom{2^{n-i}}{k} h(u)^{2^{n-i}-k} b_0^k \right) u^{2^{n+1-i}-k} \varpi^{(r+1)(i+k)}
\end{aligned}
$$

To complete the inductive step, it therefore suffices to show that whenever $i + k \leq n+1$, we have $2^{n+1-i} - k \geq 2^{n+1-i-k}$. Equivalently, and writing $j := n+1-i-k$, we must show that $2^{j+k} - k \geq 2^j$ for all $j \geq 0$, which holds as $2^k \geq k+1$ for all $k \geq 0$. □

2.3. The action of G on u. In this subsection, we study the action of G on the element $u \in W(R)_F$ corresponding to our choice of f-compatible sequence $\{\pi_n\}_n$ in $\overline{K}$ and our Frobenius lift determined by f. From the very construction of the embedding $\mathfrak{S}_F \hookrightarrow W(R)_F$ in Lemma 2.1.1, the action of $G_{\underline{\pi}}$ on u is trivial. However, for arbitrary $g \in G \setminus G_{\underline{\pi}}$, in contrast to the classical case, we know almost nothing about the shape of $g(u)$; *cf.* the discussion in §3.1. Fortunately, we are nonetheless able to prove the following facts, which are sufficient for our applications.

Define

$$I_F^{[1]} := \{x \in W(R)_F \mid \varphi^n(x) \in \mathrm{Fil}^1 W(R)_F, \forall n \geq 0\}.$$

Recall that $e_F := [F : \mathbb{Q}_p]$, and for $x \in W(R)_F$ write $\bar{x} := x \bmod \varpi \in R$. Thanks to Example 3.3.2, there exists $\mathfrak{t}_F \in W(R)_F$ satisfying $\varphi(\mathfrak{t}_F) = E\mathfrak{t}_F$. As $E \in \mathrm{Fil}^1 W(R)_F$, it is easy to see that $\varphi(\mathfrak{t}_F) \in I^{[1]}W(R)_F$, and since $\bar{\mathfrak{t}}_F^p = u^{e_0}\bar{\mathfrak{t}}_F$, we have $v_R(\overline{\varphi(\mathfrak{t}_F)}) = \frac{p}{e_F(p-1)}$.

Lemma 2.3.1. *The ideal $I_F^{[1]}$ is principal. Moreover, $x \in I_F^{[1]}$ is a generator of $I_F^{[1]}$ if and only if $v_R(\bar{x}) = \frac{p}{e_F(p-1)}$.*

Proof. When $F = \mathbb{Q}_p$, this follows immediately from [Fon94a, Proposition 5.1.3] with $r = 1$. The general case follows from a slight modification of this argument, as follows: For $y \in I_F^{[1]}$, we first claim that $v_R(\bar{y}) \geq \frac{p}{e_F(p-1)}$. To see this, we write $y = \sum_{n=0}^{\infty} \varpi^n [y_i]$ with $y_i \in R$ given by the p-power compatible sequence $y_i = (\alpha_i^{(n)})_{n\geq 0}$ for $\alpha_i^{(n)} \in \mathcal{O}_{\mathbb{C}_K}$. Then

$$0 = \theta_F(\varphi^m(y)) = \sum_{n=0}^{\infty} \varpi^n (\alpha_i^{(0)})^{p^m}.$$

By induction on n and m, it is not difficult to show that

$$v(\alpha_i^{(0)}) \geq \frac{1}{e_F} p^{-i}(1 + p^{-1} \cdots + p^{-j})$$

for all $j \geq 0$. In particular, $v_R(\bar{y}) = v(\alpha_0^{(0)}) \geq \frac{p}{e_F(p-1)}$.

Now pick a $x \in I_F^{[1]}$ with $v_R(\bar{x}) = \frac{p}{e_F(p-1)}$ (take, for example, $x = \varphi(\mathfrak{t}_F)$). Since $v_R(y) \geq v_R(x)$, we may write $y = ax + \varpi z$ with $a, z \in W(R)_F$. One checks that $z \in I_F^{[1]}$ and hence that $z \in (\varpi, x)$. An easy induction argument then shows that $y = \sum_{n=0}^{\infty} \varpi^n a_n x$, and it follows that $I_F^{[1]}$ is generated by x. □

It follows at once from Lemma 2.3.1 that $\varphi(\mathfrak{t}_F)$ is a generator of $I_F^{[1]}$. Write I^+ for the kernel of the canonical projection $\rho : W(R)_F \to W(\bar{k})_F$ induced by the projection $R \to \bar{k}$. Using the very construction of u, one checks that $u \in I^+$: Indeed, writing $\mathfrak{u} = [\underline{\pi}]$ as before, we obviously have $\mathfrak{u} \in I^+$, and it follows from the proof of Lemma 2.1.1 that $u = \lim_{n\to\infty} f^{(n)} \circ \varphi^{-n}(\mathfrak{u})$ lies in I^+ as well.

Lemma 2.3.2. *Let $g \in G$ be arbitrary. Then $g(u) - u$ lies in $I^{[1]}W(R)_F$. Moreover, if $\varpi^2 | a_1$ in $\mathcal{O}_F$ then $\frac{g(u)-u}{\varphi(\mathfrak{t}_F)}$ lies in I^+.*

Proof. As before, writing $f^{(n)} = f \circ \cdots \circ f$ for the n-fold composition of f with itself, we have $\theta_F(\varphi^n(u)) = f^{(n)}(\pi) \in K$, from which it follows that $g(u) - u$ is in $I_F^{[1]}$. By Lemma 2.3.1, we conclude that $z := \frac{g(u)-u}{\varphi(\mathfrak{t}_F)}$ lies in $W(R)_F$. It remains to show that $z \in I^+$ when $\varpi^2 | a_1$. We first observe that

$$\varphi(z) = \frac{f((g(u)) - f(u)}{\varphi^2(\mathfrak{t}_F)} = \frac{\sum_{i=1}^{p} a_i \left((g(u))^i - u^i\right)}{\varphi(E)\varphi(\mathfrak{t}_F)}.$$

For each i, we may write $(g(u))^i - u^i = (g(u) - u)h_i(g(u), u) = \varphi(\mathfrak{t}_F) z h_i(g(u), u)$ for some bivariate degree $i - 1$ homogeneous polynomials h_i with coefficients in $W(R)_F$. We therefore have

$$\varphi(E)\varphi(z) = \sum_{i=1}^{p} a_i \left(z h_i(g(u), u)\right). \tag{2.3.1}$$

Reducing modulo I^+ and noting that both u and $g(u)$ lie in I^+, we conclude from (2.3.1) that $\varpi\varphi(\rho(z)) = a_1\rho(z)$, where $\rho : W(R)_F \to W(\bar{k})_F$ is the natural projection as above. Using the fact that $v(\varphi(\rho(z))) = v(\rho(z))$, our assumption that $v(a_1) > v(\varpi)$ then implies that $\rho(z) = 0$. That is, $z \in I^+$ as desired. □

Example 2.3.3. The following example shows that the condition $\varpi^2|a_1$ in $\mathcal{O}_F$ is genuinely necessary for the conclusion of Lemma 2.3.2 to hold. Recall the situation of Example 2.1.2, with $K = \mathbb{Q}_p(\zeta_p)$, $\pi = \zeta_p - 1$, $f(u) = (u+1)^p - 1$ and $u = [\underline{\epsilon}_1] - 1$, where $\underline{\epsilon}_1 = (\zeta_{p^n})_{n\geq 1} \in R$. We may choose $g \in G$ with $g(\underline{\epsilon}_1) = \underline{\epsilon}_1^{1+p}$. We then have $g(u) - u = [\underline{\epsilon}_1]([\underline{\epsilon}_1]^p - 1)$. Now it is well-known that $[\underline{\epsilon}_1]^p - 1$ is a generator of $I^{[1]}_{\mathbb{Q}_p}$ (or one can appeal to Lemma 2.3.1). Then $z = (g(u) - u)/\varphi(\mathfrak{t}_F)$ is a unit in $W(R)$ and does not lie in I^+.

We conclude this discussion with the following lemma, needed in §5.1:

Lemma 2.3.4. *The ideal $\mathfrak{t}_F I^+ \subset W(R)_F$ is stable under the canonical action of G: that is, $g(\mathfrak{t}_F I^+) \subset \mathfrak{t}_F I^+$ for all $g \in G$.*

Proof. It is clear that I^+ is G-stable, so it suffices to show that $g(\mathfrak{t}_F) = x\mathfrak{t}_F$ for some $x \in W(R)_F$. Since $\varphi(\mathfrak{t}_F)$ is a generator of $I^{[1]}$, which is obviously G-stable from the definition, we see that $g(\varphi(\mathfrak{t}_F)) = y\varphi(\mathfrak{t}_F)$ with $y \in W(R)_F$. Hence $g(\mathfrak{t}_F) = \varphi^{-1}(y)\mathfrak{t}_F$. □

3. Étale φ-modules and Kisin modules

In this section, following Fontaine, we establish a classification of $G_{\underline{\pi}}$-representations by étale φ-modules and Kisin modules. To do this, we must first show that $K_{\underline{\pi}}/K$ is *strictly Arithmetically Profinite, or APF,* in the sense of Fontaine–Wintenberger [Win83], so that the theory of norm fields applies.

3.1. Arithmetic of f-iterate extensions.

We keep the notation and conventions of §2. Recall that our choice of an f-compatible sequence $\{\pi_n\}_n$ (in the sense that $f(\pi_n) = \pi_{n-1}$ with $\pi_0 = \pi$ a uniformizer of K) determines an element $\underline{\pi} := \{\pi_n \bmod \varpi\}_n$ of R. It also determines an infinite, totally wildly ramified extension $K_{\underline{\pi}} := \cup_{n\geq 1} K(\pi_n)$ of K, and we write $G_{\underline{\pi}} = \mathrm{Gal}(\overline{K}/K_{\underline{\pi}})$.

Lemma 3.1.1. *The extension $K_{\underline{\pi}}/K$ is strictly APF in the sense of* [Win83]*; in particular, the associated norm field $\mathbf{E}_{K_{\underline{\pi}}/K}$ is canonically identified with the subfield $k(\!(\underline{\pi})\!)$ of* $\mathrm{Fr}(R)$.

Proof. That $K_{\underline{\pi}}/K$ is strictly APF follows immediately from [CD15], which handles a more general situation. In the present setting with $f(u) \equiv u^p \bmod \varpi$, we can give a short proof as follows. As before, let us write

$$f(u) = a_1 u + a_2 u^2 + \cdots + a_{p-1}u^{p-1} + a_p u^p,$$

with $a_i \in \varpi\mathcal{O}_F$ for $1 \leq i \leq p-1$ and $a_p := 1$. For each $n \geq 1$, set $f_n := f - \pi_{n-1}$ and put $K_n := K(\pi_{n-1})$. We compute the "ramification polynomial"

$$g_n := \frac{f_n(\pi_n u + \pi_n)}{u} = \sum_{i=0}^{p-1} b_i u^i,$$

with coefficients b_i given by

$$b_i = \sum_{j=i+1}^{p} a_j \pi_n^j \binom{j}{i+1} \qquad \text{for} \qquad 0 \le i \le p-1.$$

For ease of notation, put $v_n := v_{K_{n+1}}$, and denote by $e_n := v_n(\varpi)$ the ramification index of K_{n+1}/F and by $e := v_F(p)$ the absolute ramification index of F. Since K_{n+1}/K_n is totally ramified of degree p, we have $e_n = p^n e_0$; in particular, $v_n(a_j \binom{j}{i+1} \pi_n^j) \equiv j \bmod p^n$. It follows that $v_n(b_{p-1}) = p$, and for $0 \le i \le p-2$ we have

$$v_n(b_i) = \min\{e_n e + p,\ e_n v_F(a_j) + j \ :\ i+1 \le j \le p-1\}.$$

It is easy to see that for $n \ge 1$ the lower convex hull of these points is the straight line with endpoints $(0, v_n(b_0))$ and $(p-1, p)$. In other words, defining

$$i_{\min} := \min_i \{i \ :\ \mathrm{ord}_{\varpi}(a_i) \le e,\ 1 \le i \le p\}. \tag{3.1.1}$$

the Newton polygon of g_n is a single line segment with slope the negative of

$$i_n := \frac{e_n \left(v_F(a_{i_{\min}}) + \lfloor i_{\min}/p \rfloor e\right) + i_{\min} - p}{p-1}. \tag{3.1.2}$$

In particular, for $n \ge 1$ the extension K_{n+1}/K_n is *elementary of level* i_n in the sense of [Win83, 1.3.1]; concretely, this condition means that

$$v_n(\pi_n - \sigma\pi_n) = i_n + 1 \tag{3.1.3}$$

for *every* K_n-embedding $\sigma : K_{n+1} \hookrightarrow \overline{K}$. It follows from this and [Win83, 1.4.2] that $K_{\underline{\pi}}/K$ is APF. Now let $c(K_{\underline{\pi}}/K)$ be the constant defined in [Win83, 1.2.1]. Then by [Win83, §1.4]

$$c(K_{\underline{\pi}}/K) = \inf_{n>0} \frac{i_n}{[K_{n+1} : K]},$$

so from (3.1.1) we deduce

$$\begin{aligned} c(K_{\underline{\pi}}/K) &= \inf_{n>0} \frac{e_n \left(v_F(a_{i_{\min}}) + \lfloor i_{\min}/p \rfloor e\right) + i_{\min} - p}{p^n(p-1)} \\ &= \frac{e_0}{p-1}\left(v_F(a_{i_{\min}}) + \lfloor i_{\min}/p \rfloor e\right) - \frac{p - i_{\min}}{p(p-1)} \end{aligned}$$

since $p - i_{\min} \ge 0$, so the above infimum occurs when $n = 1$. As $i_{\min} \ge 1$, the above constant is visibly positive, so by the very definition [Win83, 1.2.1], $K_{\underline{\pi}}/K$ is *strictly* APF.

The canonical embedding of the norm field of $K_{\underline{\pi}}/K$ into $\mathrm{Fr}(R)$ is described in [Win83, §4.2]; that the image of this embedding coincides with $k(\!(\underline{\pi})\!)$ is a consequence of [Win83, 2.2.4, 2.3.1]. $\square$

Remark 3.1.2. Observe that if the coefficient a_1 of the linear term of $f(u)$ has $v(a_1) \le 1$, then we have $i_{\min} = 1$ and

$$c(K_{\underline{\pi}}/K) = \frac{e_0}{p-1} v_F(a_1) - \frac{1}{p}.$$

In this situation, $v_F(a_1)$—which plays an important role in our theory—is encoded in the ramification structure of $K_{\underline{\pi}}/K$.

It is natural to ask when two given polynomials f and f' with corresponding compatible choices $\underline{\pi}$ and $\underline{\pi}'$ give rise to the same iterate extension. Let us write $f(x) = x^p + a_{p-1}x^{p-1} + \cdots + a_1 x$ and $f'(x) = x^p + a'_{p-1}x^{p-1} + \cdots + a'_1 x$, with $a_i, a'_i \in \mathcal{O}_F$ and $a_i \equiv a'_i \equiv 0 \bmod \varpi$ for $1 \le i < p$. Let $\{\pi_n\}$ (respectively $\{\pi'_n\}$) be an f (resp. f') compatible sequence of elements in $\overline{K}$. Set $K_n := K(\pi_{n-1})$ (resp. $K'_n = K(\pi'_{n-1})$) and let $a_s u^s$ and $a'_{s'}u^{s'}$ be the lowest degree terms of $f(u)$ and $f'(u)$ respectively.

Proposition 3.1.3. *If $K_{\underline{\pi}} = K_{\underline{\pi}'}$ as subfields of $\overline{K}$, then $K_n = K'_n$ for all $n \ge 1$ and there exists an invertible power series $\xi(x) \in \mathcal{O}_F[\![x]\!]$ with $\xi(x) = \mu_0 x + \cdots$ and $\mu_0 \in \mathcal{O}_F^\times$ such that*

$$f(\xi(x)) = \xi(f'(x)).$$

In particular, $s = s'$ and $v(a_s) = v(a'_s)$ are numerical invariants of $K_{\underline{\pi}} = K_{\underline{\pi}'}$. Conversely, if f and f' are given with $s = s'$ and $v(a_s) = v(a'_s)$, then we have $a_s = \mu_0^{1-s}a'_s$ for a unique $\mu_0 \in \mathcal{O}_F^\times$ and there is a unique power series $\xi(x) \in F[\![x]\!]$ with $\xi(x) \equiv \mu_0 x \bmod x^2$ satisfying $f(\xi(x)) = \xi(f'(x))$ as formal power series in $F[\![x]\!]$. If $\xi(x)$ lies in $\mathcal{O}_F[\![x]\!]$, then for any choice $\{\pi'_n\}_n$ of f'-compatible sequence with π'_0 a uniformizer of K, the sequence defined by $\pi_n := \xi(\underline{\pi}'_n)$ is f-compatible with $\pi_0 = \xi(\pi'_0)$ a uniformizer of K and $K_{\underline{\pi}} = K_{\underline{\pi}'}$. Furthermore, if $v(a_s) = v(a'_s) = v(\varpi)$, then $\xi(x)$ always lies in $\mathcal{O}_F[\![x]\!]$.

Proof. Suppose first that $K_{\underline{\pi}} = K_{\underline{\pi}'}$, and write simply K_∞ for this common, strictly APF extension of K in $\overline{K}$. It follows from the proof of Lemma 3.1.1 that K_{n+1} and K'_{n+1} are both the n-th elementary subextension of K_∞; *i.e.* the fixed field of $G_K^{b_n}G_{K_\infty}$, where b_n is the n-th break in the ramification filtration $G_K^u G_{K_\infty}$; see [Win83, 1.4]. In particular, $K_{n+1} = K'_{n+1}$ for all $n \ge 0$. Now let $W_\varpi(\bullet)$ be the functor of ϖ-Witt vectors; it is the unique functor from $\mathcal{O}_F$-algebras to $\mathcal{O}_F$-algebras satisfying

(1) For any $\mathcal{O}_F$-algebra A, we have $W_\varpi(\bullet) = \prod_{n\ge 0}\bullet =: \bullet^{\mathbf{N}}$ as functors from $\mathcal{O}_F$-algebras to sets.
(2) The ghost map $W_\varpi(\bullet) \to \bullet^{\mathbf{N}}$ given by

$$(a_0, a_1, a_2, \ldots) \mapsto (a_0, a_0^p + \varpi a_1, a_0^{p^2} + \varpi a_1^p + \varpi^2 a_2, \ldots)$$

is a natural transformation of functors from $\mathcal{O}_F$-algebras to $\mathcal{O}_F$-algebras.

We remark that $W_\varpi(\bullet)$ exists and depends only on ϖ, and is equipped with a unique natural transformation $\varphi : W_\varpi(\bullet) \to W_\varpi(\bullet)$ which on ghost components has the effect $(a_0, a_1, \ldots) \mapsto (a_1, a_2, \ldots)$; see [CD15, §2].
Define the ring

$$\mathbf{A}^+_{K_\infty/K} := \{(\underline{x}_i)_i \in \varprojlim_{\varphi} W_\varpi(\mathcal{O}_{\widehat{K}_\infty}) \ : \ \underline{x}_n \in W_\varpi(\mathcal{O}_{K_{n+1}}) \text{ for all } n\},$$

which depends only on F, ϖ, and K_∞/K. The main theorem of [CD15], implies that $\mathbf{A}^+_{K_\infty/K}$ is a ϖ-adically complete and separated $\mathcal{O}_F$-algebra equipped with a Frobenius endomorphism φ, which is canonically a Frobenius-stable subring of $W(R)_F$ that is closed under the weak topology on $W(R)_F$. Giving $\mathbf{A}^+_{K_\infty/K}$ the subspace topology, the proof of [CD15, Prop. 7.13] then shows that the f (respectively f')-compatible sequence $\underline{\pi}$ (respectively $\underline{\pi}'$) determine isomorphisms of topological $\mathcal{O}_F$-algebras

$$\eta, \eta' : \mathcal{O}_F[\![x]\!] \longrightarrow \mathbf{A}^+_{K_\infty/K}$$

characterized by the requirement that the ghost components of $(\eta)_n$ (resp. $(\eta')_n$) are $(\pi_n, f(\pi_n), f^{(2)}(\pi_n), \ldots)$ (resp. $(\pi'_n, f'(\pi'_n), f'^{(2)}(\pi'_n), \ldots)$); here we give $\mathcal{O}_F[\![x]\!]$ the (ϖ, x)-adic topology. These isomorphisms moreover satisfy

$$\eta(f(x)) = \varphi(\eta(x)) \qquad \text{and} \qquad \eta'(f'(x)) = \varphi(\eta'(x)).$$

We therefore obtain a continuous automorphism $\xi : \mathcal{O}_F[\![x]\!] \to \mathcal{O}_F[\![x]\!]$ satisfying

$$f(\xi(x)) = \xi(f'(x)). \tag{3.1.4}$$

Since ξ is a continuous automorphism of $\mathcal{O}_F[\![x]\!]$, we have that ξ preserves the maximal ideal (ϖ, x). This implies that $\xi(x) \equiv \mu_0 x \bmod x^2$ with $\mu_0 \in \mathcal{O}_F^\times$. Then (3.1.4) forces $a_s\mu_0^s x^s = a'_{s'}\mu_0 x^{s'}$ which implies $s = s'$ and $v(a_s) = v(a'_s)$. Conversely, suppose given f and f' with $s = s'$ and $v(a_s) = v(a'_s)$ and let $\mu_0 \in \mathcal{O}_F^\times$ be the unique unit with $a_s = \mu_0^{1-s}a'_s$; note that this exists because $s - 1 < p$. We inductively construct degree i polynomials $\xi_i(x) = \sum_{j=1}^i \mu_j x^j$ so that $f(\xi_i(x)) \equiv \xi_i(f'(x)) \bmod x^{i+s}$. As $\mu_0^s a_s = \mu_0 a'_s$, we may clearly take $\xi_1(x) = \mu_0 x$. If $\xi_i(x)$ has been constructed, we write $\xi_{i+1}(x) = \xi_i(x) + \mu_{i+1}x^{i+1}$ and $f(\xi_i(x)) - \xi_i(f'(x)) \equiv \lambda x^{i+s} \bmod x^{i+s+1}$ and seek to solve

$$f(\xi_{i+1}(x)) \equiv \xi_{i+1}(f'(x)) \bmod x^{i+s+1}. \tag{3.1.5}$$

As $f(\xi_{i+1}(x)) = f(\xi_i(x)) + \frac{df}{dx}(\xi_i(x))(\mu_{i+1}x^{i+1}) + \cdots$, we see that (3.1.5) is equivalent to

$$\lambda = \mu_{i+1}(a_1 - {a'_1}^{i+1}) \text{ if } s = 1, \quad \text{and} \quad \lambda = \mu_{i+1} s a_s \mu_0^{s-1} \text{ if } s > 1 \tag{3.1.6}$$

which admits a unique solution $\mu_{i+1} \in F$. We set $\xi(x) = \lim_i \xi_i(x) \in F[\![x]\!]$, which by construction satisfies the desired intertwining relation (3.1.4). If ξ lies in $\mathcal{O}_F[\![x]\!]$, it is clear that any f'-compatible sequence π'_n with π'_0 a uniformizer of K yields an f-compatible sequence $\pi_n := \xi(\pi'_n)$ with π_0 a uniformizer of K and $K_n := K(\pi_{n-1}) = K(\pi'_{n-1}) = K'_n$ for all $n \geq 1$. Finally, since we have $f(x) = f'(x) \equiv x^p \bmod \varpi$, it follows that $f(\xi_i(x)) - \xi_i(f'(x)) \equiv 0 \bmod \varpi$, *i.e.* $\lambda \equiv 0 \bmod \varpi$ in the above construction. When $v(a_s) = v(a'_s) = v(\varpi)$, it then follows from (3.1.6) that $\mu_{i+1} \in \mathcal{O}_F$, and $\xi(x) \in \mathcal{O}_F[\![x]\!]$ as claimed. □

As an immediate consequence of Proposition (3.1.3), one sees that there are infinitely many distinct f-iterate extensions $K_{\underline{\pi}}$ inside of $\overline{K}$.

3.2. Étale φ-modules. Let $\mathcal{O}_{\mathcal{E}}$ be the p-adic completion of $\mathfrak{S}_F[1/u]$, equipped with the unique continuous extension of φ. Our fixed embedding $\mathfrak{S}_F \hookrightarrow W(R)$ determined by f and $\underline{\pi}$ uniquely extends to a φ-equivariant embedding $\iota : \mathcal{O}_{\mathcal{E}} \hookrightarrow W(\operatorname{Fr} R)_F$, and we identify $\mathcal{O}_{\mathcal{E}}$ with its image in $W(\operatorname{Fr} R)_F$. We note that $\mathcal{O}_{\mathcal{E}}$ is a complete discrete valuation ring with uniformizer ϖ and residue field $k(\!(\underline{\pi})\!)$, which, as a subfield of $\operatorname{Fr} R$, coincides with the norm field of $K_{\underline{\pi}}/K$ thanks to Lemma 3.1.1. As $\operatorname{Fr} R$ is algebraically closed, there is a unique separable closure $k(\!(\underline{\pi})\!)^{\mathrm{sep}}$ of $k(\!(\underline{\pi})\!)$ in $\operatorname{Fr} R$, and the maximal unramified extension (*i.e.* strict Henselization) $\mathcal{O}_{\mathcal{E}^{\mathrm{ur}}}$ of $\mathcal{O}_{\mathcal{E}}$ with residue field $k(\!(\underline{\pi})\!)^{\mathrm{sep}}$ is uniquely determined up to unique isomorphism. The universal property of strict Henselization guarantees that ι uniquely extends to an embedding $\mathcal{O}_{\mathcal{E}^{\mathrm{ur}}} \hookrightarrow W(\operatorname{Fr} R)_F$, which moreover realizes $\mathcal{O}_{\mathcal{E}^{\mathrm{ur}}}$ as a φ-stable subring. We write $\mathcal{O}_{\widehat{\mathcal{E}}^{\mathrm{ur}}}$ for the p-adic completion of $\mathcal{O}_{\mathcal{E}^{\mathrm{ur}}}$, which is again a φ-stable subring of $W(\operatorname{Fr} R)_F$. Again using the universal property of strict Henselization, one sees that each of $\mathcal{O}_{\mathcal{E}}, \mathcal{O}_{\mathcal{E}^{\mathrm{ur}}}$ and $\mathcal{O}_{\widehat{\mathcal{E}}^{\mathrm{ur}}}$ are $G_{\underline{\pi}}$-stable subrings of $W(\operatorname{Fr} R)_F$, with $G_{\underline{\pi}}$ acting trivially on $\mathcal{O}_{\mathcal{E}}$. As suggested by the notation, we write $\mathcal{E}$, $\mathcal{E}^{\mathrm{ur}}$, and $\widehat{\mathcal{E}}^{\mathrm{ur}}$ for the fraction fields of $\mathcal{O}_{\mathcal{E}}, \mathcal{O}_{\mathcal{E}^{\mathrm{ur}}}$ and $\mathcal{O}_{\widehat{\mathcal{E}}^{\mathrm{ur}}}$, respectively. Finally, we define $\mathfrak{S}_F^{\mathrm{ur}} := W(R)_F \cap \mathcal{O}_{\widehat{\mathcal{E}}^{\mathrm{ur}}}$.

Lemma 3.2.1. *With notation as above:*

(1) *The natural action of $G_{\underline{\pi}}$ on $\mathcal{O}_{\widehat{\mathcal{E}}^{\mathrm{ur}}}$ induces an isomorphism of profinite groups*

$$G_{\underline{\pi}} := \operatorname{Gal}(\overline{K}/K_{\underline{\pi}}) \simeq \operatorname{Aut}(\mathcal{O}_{\widehat{\mathcal{E}}^{\mathrm{ur}}}/\mathcal{O}_{\mathcal{E}}) = \operatorname{Gal}(\widehat{\mathcal{E}}^{\mathrm{ur}}/\mathcal{E}).$$

(2) *The inclusions $\mathcal{O}_F \hookrightarrow (\mathcal{O}_{\widehat{\mathcal{E}}^{\mathrm{ur}}})^{\varphi=1}$ and $\mathcal{O}_{\mathcal{E}} \hookrightarrow (\mathcal{O}_{\widehat{\mathcal{E}}^{\mathrm{ur}}})^{G_{\underline{\pi}}}$ are isomorphisms.*

Proof. By the very construction of $\mathcal{O}_{\widehat{\mathcal{E}}^{\mathrm{ur}}}$ and the fact that the residue field of $\mathcal{O}_{\mathcal{E}}$ is identified with the norm field $\mathbf{E}_{K_{\underline{\pi}}/K}$ by Lemma 3.1.1, we have an isomorphism of topological groups $\operatorname{Gal}(\mathbf{E}^{\mathrm{sep}}_{K_{\underline{\pi}}/K}/\mathbf{E}_{K_{\underline{\pi}}/K}) \simeq \operatorname{Aut}(\mathcal{O}_{\widehat{\mathcal{E}}^{\mathrm{ur}}}/\mathcal{O}_{\mathcal{E}})$ by the theory of unramified extensions of local fields. On the other hand, the theory of norm fields [Win83, 3.2.2] provides a natural isomorphism of topological groups $G_{\underline{\pi}} \simeq \operatorname{Gal}(\mathbf{E}^{\mathrm{sep}}_{K_{\underline{\pi}}/K}/\mathbf{E}_{K_{\underline{\pi}}/K})$, giving (1).

To prove (2), note that the maps in question are local maps of ϖ-adically separated and complete local rings, so by a standard successive approximation argument it suffices to prove that these maps are surjective modulo ϖ. Now left-exactness of φ-invariants (respectively $G_{\underline{\pi}}$-invariants) gives an $\mathbf{F}_p$-linear (respectively $\mathbf{E}_{K_{\underline{\pi}}/K}$-linear) injection

$$(\mathcal{O}_{\widehat{\mathcal{E}}^{\mathrm{ur}}})^{\varphi=1}/(\varpi) \hookrightarrow (\mathbf{E}^{\mathrm{sep}}_{K_{\underline{\pi}}/K})^{\varphi=1} = \mathbf{F}_p = \mathcal{O}_F/(\varpi),$$

respectively

$$(\mathcal{O}_{\widehat{\mathcal{E}}^{\mathrm{ur}}})^{G_{\underline{\pi}}}/(\varpi) \hookrightarrow (\mathbf{E}^{\mathrm{sep}}_{K_{\underline{\pi}}/K})^{G_{\underline{\pi}}} = \mathbf{E}_{K_{\underline{\pi}}/K} = \mathcal{O}_{\mathcal{E}}/(\varpi)$$

which must be an isomorphism of vector spaces over $\mathbf{F}_p$ (respectively $\mathbf{E}_{K_{\underline{\pi}/K}}$) as the source is nonzero and the target is 1-dimensional. We conclude that $\mathcal{O}_F \hookrightarrow (\mathcal{O}_{\widehat{\mathcal{E}}^{\mathrm{ur}}})^{\varphi=1}$ (respectively $\mathcal{O}_{\mathcal{E}} \hookrightarrow (\mathcal{O}_{\widehat{\mathcal{E}}^{\mathrm{ur}}})^{G_{\underline{\pi}}}$) is surjective modulo ϖ, and therefore an isomorphism as desired. □

Denote by $\mathrm{Mod}^{\varphi}_{\mathcal{O}_{\mathcal{E}}}$ (resp. $\mathrm{Mod}^{\varphi,\mathrm{tor}}_{\mathcal{O}_{\mathcal{E}}}$) the category of pairs (M, φ_M), where M is a finite free $\mathcal{O}_{\mathcal{E}}$-module (resp. a finite $\mathcal{O}_{\mathcal{E}}$-module killed by a power of ϖ) and $\varphi_M : M \to M$ is a φ-semilinear and additive map whose linearization $1 \otimes \varphi_M : \varphi^* M \to M$ is an isomorphism. In each case, morphisms are φ-equivarant $\mathcal{O}_{\mathcal{E}}$-module homomorphisms. Let $\mathrm{Rep}_{\mathcal{O}_F}(G_{\underline{\pi}})$ (resp. $\mathrm{Rep}^{\mathrm{tor}}_{\mathcal{O}_F}(G_{\underline{\pi}})$) be the category of finite, free $\mathcal{O}_F$-modules (resp. finite $\mathcal{O}_F$-modules killed by a power of ϖ) that are equipped with a continuous and $\mathcal{O}_F$-linear action of $G_{\underline{\pi}}$. For M in $\mathrm{Mod}^{\varphi}_{\mathcal{O}_{\mathcal{E}}}$ or in $\mathrm{Mod}^{\varphi,\mathrm{tor}}_{\mathcal{O}_{\mathcal{E}}}$, we define

$$\underline{V}(M) := (\mathcal{O}_{\widehat{\mathcal{E}}^{\mathrm{ur}}} \otimes_{\mathcal{O}_{\mathcal{E}}} M)^{\varphi=1},$$

which is an $\mathcal{O}_F$-module with a continuous action of $G_{\underline{\pi}}$. For V in $\mathrm{Rep}_{\mathcal{O}_F}(G_{\underline{\pi}})$ or in $\mathrm{Rep}^{\mathrm{tor}}_{\mathcal{O}_F}(G_{\underline{\pi}})$, we define

$$\underline{M}(V) = (\mathcal{O}_{\widehat{\mathcal{E}}^{\mathrm{ur}}} \otimes_{\mathcal{O}_F} V)^{G_{\underline{\pi}}},$$

which is an $\mathcal{O}_{\mathcal{E}}$-module with a φ-semilinear endomorphism $\varphi_{\underline{M}} := \varphi_{\mathcal{O}_{\widehat{\mathcal{E}}^{\mathrm{ur}}}} \otimes 1$.

Theorem 3.2.2. *The functors $\underline{V}$ and $\underline{M}$ are quasi-inverse equivalences between the exact tensor categories* $\mathrm{Mod}^{\varphi}_{\mathcal{O}_{\mathcal{E}}}$ (*resp.* $\mathrm{Mod}^{\varphi,\mathrm{tor}}_{\mathcal{O}_{\mathcal{E}}}$) *and* $\mathrm{Rep}_{\mathcal{O}_F}(G_{\underline{\pi}})$ (*resp.* $\mathrm{Rep}^{\mathrm{tor}}_{\mathcal{O}_F}(G_{\underline{\pi}})$).

Proof. As in the proof of [KR09, Theorem 1.6], the original arguments of Fontaine [Fon90, A1.2.6] carry over to the present situation. Indeed, by standard arguments with inverse limits, it is enough to prove the Theorem for ϖ-power torsion objects. To do so, one first proves that $\underline{M}$ is exact, which by (faithful) flatness of the inclusion $\mathcal{O}_{\mathcal{E}} \hookrightarrow \mathcal{O}_{\mathcal{E}^{\mathrm{ur}}}$ amounts to the vanishing of $H^1(G_{\underline{\pi}}, \cdot)$ on the category of finite length $\mathcal{O}_{\mathcal{E}^{\mathrm{ur}}}$-modules with a continuous semilinear $G_{\underline{\pi}}$-action. By a standard dévissage, such vanishing is reduced to the case of modules killed by ϖ, where it follows from Hilbert's Theorem 90 and Lemma 3.2.1. One then checks that for any torsion V, the natural comparison map $\underline{M}(V) \otimes_{\mathcal{O}_{\mathcal{E}}} \mathcal{O}_{\mathcal{E}^{\mathrm{ur}}} \to V \otimes_{\mathcal{O}_F} \mathcal{O}_{\mathcal{E}^{\mathrm{ur}}}$ induced by multiplication in $\mathcal{O}_{\mathcal{E}^{\mathrm{ur}}}$ is an $\mathcal{O}_{\mathcal{E}^{\mathrm{ur}}}$-linear, φ, and $G_{\underline{\pi}}$-compatible isomorphism by dévissage (using the settled exactness of $\underline{M}$) to the case that V is ϖ-torsion, where it again follows from Hilbert Theorem 90. Passing to submodules on which φ acts as the identity and using Lemma 3.2.1(2) then gives a natural isomorphism $\underline{V} \circ \underline{M} \simeq \mathrm{id}$.

In a similar fashion, the exactness of $\underline{V}$ and the fact that the comparison map

$$\underline{V}(M) \otimes_{\mathcal{O}_F} \mathcal{O}_{\mathcal{E}^{\mathrm{ur}}} \longrightarrow M \otimes_{\mathcal{O}_{\mathcal{E}}} \mathcal{O}_{\mathcal{E}^{\mathrm{ur}}} \tag{3.2.1}$$

induced by multiplication in $\mathcal{O}_{\mathcal{E}^{\mathrm{ur}}}$ is an isomorphism for general ϖ-power torsion modules M follows by dévissage from the the truth of these claims in the case of M killed by ϖ. In this situation, the comparison map (3.2.1) is shown to be injective by checking that any $\mathbf{F}_p$-linearly independent set of vectors in

$\underline{V}(M)$ remains $\mathbf{E}^{\mathrm{sep}}_{K_\infty/K}$-linearly independent in $\mathbf{E}^{\mathrm{sep}}_{K_\infty/K} \otimes_{\mathbf{F}_p} \underline{V}(M)$, which is accomplished by a standard argument using the Frobenius endomorphism and Lemma 3.2.1(2). To check surjectivity is then a matter of showing that both sides of (3.2.1) have the same $\mathbf{E}^{\mathrm{sep}}_{K_\infty/K}$-dimension, *i.e.* that the $\mathbf{F}_p$-vector space $\underline{V}(M)$ has dimension $d := \dim_{\mathbf{E}_{K_\infty/K}} M$. Equivalently, we must prove that $\underline{V}(M)$ has p^d elements. Identifying M with $\mathbf{E}^d_{K_\infty/K}$ by a choice of $\mathbf{E}_{K_\infty/K}$-basis and writing (c_{ij}) for the resulting matrix of φ, one (noncanonically) realizes $\underline{V}(M)$ as the set of $\mathbf{E}^{\mathrm{sep}}_{K_\infty/K}$-solutions to the system of d-equations $x_i^p = \sum c_{ij} x_j$ in d-unknowns, which has exactly p^d solutions as φ is étale, so the matrix (c_{ij}) is *invertible.* □

In what follows, we will need a contravariant version of Theorem 3.2.2, which follows from it by a standard duality argument (*e.g.* [Fon90, §1.2.7]). For any $M \in \mathrm{Mod}^{\varphi}_{\mathcal{O}_\mathcal{E}}$ (respectively $M \in \mathrm{Mod}^{\varphi,\mathrm{tor}}_{\mathcal{O}_\mathcal{E}}$), we define

$$\underline{T}(M) := \mathrm{Hom}_{\mathcal{O}_\mathcal{E},\varphi}(M, \mathcal{O}_{\widehat{\mathcal{E}}^{\mathrm{ur}}}), \quad \text{respectively} \quad \underline{T}(M) := \mathrm{Hom}_{\mathcal{O}_\mathcal{E},\varphi}(M, \widehat{\mathcal{E}}^{\mathrm{ur}}/\mathcal{O}_{\widehat{\mathcal{E}}^{\mathrm{ur}}}),$$

which is naturally an $\mathcal{O}_F$-module with a continuous action of $G_{\underline{\pi}}$.

Corollary 3.2.3. *The contravariant functor $\underline{T}$ induces an anti-equivalence between* $\mathrm{Mod}^{\varphi}_{\mathcal{O}_\mathcal{E}}$ (*resp.* $\mathrm{Mod}^{\varphi,\mathrm{tor}}_{\mathcal{O}_\mathcal{E}}$) *and* $\mathrm{Rep}_{\mathcal{O}_F}(G_{\underline{\pi}})$ (*resp.* $\mathrm{Rep}^{\mathrm{tor}}_{\mathcal{O}_F}(G_{\underline{\pi}})$).

3.3. Kisin modules and Representations of finite E-height. For an integer $r \geq 0$, we write $'\mathrm{Mod}^{\varphi,r}_{\mathfrak{S}_F}$ for the category of finite-type $\mathfrak{S}_F$-modules $\mathfrak{M}$ equipped with a $\varphi_{\mathfrak{S}_F}$-semilinear endomorphism $\varphi_\mathfrak{M} : \mathfrak{M} \to \mathfrak{M}$ satisfying

- the cokernel of the linearization $1 \otimes \varphi : \varphi^*\mathfrak{M} \to \mathfrak{M}$ is killed by E^r;
- the natural map $\mathfrak{M} \to \mathcal{O}_\mathcal{E} \otimes_{\mathfrak{S}_F} \mathfrak{M}$ is injective.

One checks that together these conditions guarantee that the scalar extension $\mathcal{O}_\mathcal{E} \otimes_{\mathfrak{S}_F} \mathfrak{M}$ is an object of $\mathrm{Mod}^{\varphi}_{\mathcal{O}_\mathcal{E}}$ when $\mathfrak{M}$ is torsion free, and an object of $\mathrm{Mod}^{\varphi,\mathrm{tor}}_{\mathcal{O}_\mathcal{E}}$ if $\mathfrak{M}$ is killed by a power of ϖ. Morphisms in $'\mathrm{Mod}^{\varphi,r}_{\mathfrak{S}_F}$ are φ-compatible $\mathfrak{S}_F$-module homomorphisms. By definition, the category of *Kisin modules of E-height r*, denoted $\mathrm{Mod}^{\varphi,r}_{\mathfrak{S}_F}$, is the full subcategory of $'\mathrm{Mod}^{\varphi,r}_{\mathfrak{S}_F}$ consisting of those objects which are finite and free over $\mathfrak{S}_F$. For any such Kisin module $\mathfrak{M} \in \mathrm{Mod}^{\varphi,r}_{\mathfrak{S}_F}$, we define

$$T_\mathfrak{S}(\mathfrak{M}) := \mathrm{Hom}_{\mathfrak{S}_F,\varphi}(\mathfrak{M}, \mathfrak{S}^{\mathrm{ur}}_F),$$

with $\mathfrak{S}^{\mathrm{ur}}_F := W(R)_F \cap \mathcal{O}_{\widehat{\mathcal{E}}^{\mathrm{ur}}}$ as above Lemma 3.2.1; this is naturally an $\mathcal{O}_F$-module with a linear action of $G_{\underline{\pi}}$.

Proposition 3.3.1. *Let $\mathfrak{M} \in \mathrm{Mod}^{\varphi,r}_{\mathfrak{S}_F}$ and write $M = \mathcal{O}_\mathcal{E} \otimes_{\mathfrak{S}_F} \mathfrak{M}$ for the corresponding object of* $\mathrm{Mod}^{\varphi}_{\mathcal{O}_\mathcal{E}}$.

(1) *There is a canonical $\mathcal{O}_F$-linear and $G_{\underline{\pi}}$-equivariant isomorphism $T_\mathfrak{S}(\mathfrak{M}) \simeq \underline{T}(M)$. In particular, $T_\mathfrak{S}(\mathfrak{M})$ is an object of* $\mathrm{Rep}_{\mathcal{O}_F}(G_{\underline{\pi}})$ *and* $\mathrm{rank}_{\mathcal{O}_F}(T_\mathfrak{S}(\mathfrak{M})) = \mathrm{rank}_{\mathfrak{S}_F}(\mathfrak{M})$.

(2) *The inclusion $\mathfrak{S}^{\mathrm{ur}}_F \hookrightarrow W(R)_F$ induces a natural isomorphism of $\mathcal{O}_F[G_{\underline{\pi}}]$-modules $T_\mathfrak{S}(\mathfrak{M}) \simeq \mathrm{Hom}_{\mathfrak{S}_F,\varphi}(\mathfrak{M}, W(R)_F)$.*

Proof. As in the proofs of [Kis06, 2.1.2, 2.1.4] and [KR09, 3.2.1], the Lemma follows from B1.4.2 and B1.8.3 of [Fon90] (*cf.* B1.8.6), using [Fon90, A1.2] and noting that Fontaine's arguments—which are strictly speaking only for $F = \mathbb{Q}_p$—carry over *mutatis mutandis* to our more general situation. ☐

Example 3.3.2. Let $\mathfrak{M}$ be a Kisin module of rank 1 over $\mathfrak{S}_F$. Choosing a basis $\mathfrak{e}$ of $\mathfrak{M}$ and identifying $\mathfrak{M} = \mathfrak{S}_F \cdot \mathfrak{e}$, it follows from Weierstrass preparation that we must have $\varphi(\mathfrak{e}) = \mu E^m \mathfrak{e}$ for some $\mu \in \mathfrak{S}_F^\times$. Consider the particular case that $\varphi(\mathfrak{e}) = E\mathfrak{e}$, which is a rank-1 Kisin module of E-height 1. Proposition (3.3.1) then shows that $T_{\mathfrak{S}}(\mathfrak{M})$ gives an $\mathcal{O}_F$-valued character of $G_{\underline{\pi}}$ and that there exists $\mathfrak{t} \in W(R)_F$ satisfying $\varphi(\mathfrak{t}) = E\mathfrak{t}$ and $\mathfrak{t} \bmod \varpi \neq 0$ inside R. We will see in §5 that the character of $G_{\underline{\pi}}$ furnished by $T_{\mathfrak{S}}(\mathfrak{M})$ can be extended to a Lubin-Tate character of G if we assume that $\varpi^2 | a_1$ in $\mathcal{O}_F$, where a_1 is the linear coefficient of $f(x) \in \mathcal{O}_F[x]$.

Let $\mathrm{Rep}_F(G_{\underline{\pi}})$ denote the category of continuous, F-linear representations of $G_{\underline{\pi}}$. An object V of $\mathrm{Rep}_F(G_{\underline{\pi}})$ is of *E-height r* if there exists a Kisin module $\mathfrak{M} \in \mathrm{Mod}_{\mathfrak{S}_F}^{\varphi,r}$ with $V \simeq T_{\mathfrak{S}_F}(\mathfrak{M})[1/p]$, and V is of *finite E-height* if there exists an integer r such that V is of E-height r. As $E = E(u)$ is fixed throughout the paper, we will simply say that V is of (*finite*) *height r*.

For $\mathfrak{M}$ an arbitrary object of $\mathrm{Mod}_{\mathfrak{S}_F}^{\varphi,r}$, we write $V_{\mathfrak{S}}(\mathfrak{M}) := T_{\mathfrak{S}}(\mathfrak{M})[1/p]$ for the associated height-r representation of $G_{\underline{\pi}}$. We will need the following generalization of [Kis06, Lemma 2.1.15] (or [Liu07, Corollary 2.3.9]):

Proposition 3.3.3. *If $V \in \mathrm{Rep}_F(G_{\underline{\pi}})$ is of height r then for any $G_{\underline{\pi}}$-stable $\mathcal{O}_F$-lattice $L \subset V$, there exists $\mathfrak{N} \in \mathrm{Mod}_{\mathfrak{S}_F}^{\varphi,r}$ such that $T_{\mathfrak{S}}(\mathfrak{N}) \simeq L$ in $\mathrm{Rep}_{\mathcal{O}_F}(G_{\underline{\pi}})$.*

The proof of Proposition 3.3.3 will make use of the following key lemma:

Lemma 3.3.4. *Let $\mathfrak{M}$ be an object of $'\mathrm{Mod}_{\mathfrak{S}_F}^{\varphi,r}$ that is torsion-free. Then the intersection $\mathfrak{M}' := \mathfrak{M}[1/p] \cap (\mathcal{O}_{\mathcal{E}} \otimes_{\mathfrak{S}_F} \mathfrak{M})$, taken inside of $\mathcal{E} \otimes_{\mathfrak{S}_F} \mathfrak{M}$, is an object in $\mathrm{Mod}_{\mathfrak{S}_F}^{\varphi,r}$ and there are canonical inclusions $\mathfrak{M} \subset \mathfrak{M}' \subset \mathcal{O}_{\mathcal{E}} \otimes_{\mathfrak{S}_F} \mathfrak{M}$.*

Proof. The proof of Lemma 2.3.7 in [Liu07] carries over *mutatis mutandis* to the present situation. ☐

Proof of Proposition 3.3.3. As the proof is a simple adaptation of that of Corollary 2.3.9 in [Liu07], we simply sketch the highlights. Let $V \in \mathrm{Rep}_F(G_{\underline{\pi}})$ be of height r, and select $\mathfrak{M} \in \mathrm{Mod}_{\mathfrak{S}_F}^{\varphi,r}$ with $V \simeq V_{\mathfrak{S}}(\mathfrak{M})$. Put $T := T_{\mathfrak{S}}(\mathfrak{M})$, which is a $G_{\underline{\pi}}$-stable $\mathcal{O}_F$-lattice in V, and let $L \subset V$ be an arbitrary $G_{\underline{\pi}}$-stable $\mathcal{O}_F$-lattice. Put $M := \mathcal{O}_{\mathcal{E}} \otimes_{\mathfrak{S}_F} \mathfrak{M}$ and let $N \in \mathrm{Mod}_{\mathcal{O}_{\mathcal{E}}}^{\varphi}$ be the object of $\mathrm{Mod}_{\mathcal{O}_{\mathcal{E}}}^{\varphi}$ corresponding to L via Corollary 3.2.3, so $\underline{T}(N) \simeq L$ in $\mathrm{Rep}_{\mathcal{O}_F}(G_{\underline{\pi}})$. Without loss of generality, we may assume that $N \subset M$. Writing $\mathfrak{f} : M \to M/N$ for the natural projection, it is easy to check that $\mathfrak{f}(\mathfrak{M})$ is an object of $'\mathrm{Mod}_{\mathfrak{S}_F}^{\varphi,r}$. It then follows from Proposition [Fon90, B 1.3.5] that $\mathfrak{N}' := \ker(\mathfrak{f}|_{\mathfrak{M}}) \subset N$ is an object of $'\mathrm{Mod}_{\mathfrak{S}_F}^{\varphi,r}$. Writing $\mathfrak{N} := \mathfrak{N}'[1/p] \cap N$, we have that $\mathfrak{N}$ is an object of $\mathrm{Mod}_{\mathfrak{S}_F}^{\varphi,r}$ thanks to Lemma 3.3.4, and by construction we have $\mathcal{O}_{\mathcal{E}} \otimes_{\mathfrak{S}_F} \mathfrak{N} \simeq N$, so that $T_{\mathfrak{S}}(\mathfrak{N}) \simeq L$ as $\mathcal{O}_F[G_{\underline{\pi}}]$-modules thanks to Proposition 3.3.1 and the choice of N. $\square$

Proposition 3.3.5. *Assume that $\varphi^n(f(u)/u)$ is not power of $E = E(u)$ for any $n \geq 0$. Then the functor $T_{\mathfrak{S}} : \mathrm{Mod}_{\mathfrak{S}_F}^{\varphi,r} \rightsquigarrow \mathrm{Rep}_{\mathcal{O}_F}(G_{\underline{\pi}})$ is fully faithful.*

Proof. We use an idea of Caruso [Car, Proposition 3.1]. Fix $\mathfrak{M}, \mathfrak{M}' \in \mathrm{Mod}_{\mathfrak{S}_F}^{\varphi,r}$. Using Corollary 3.2.3 and Lemma 3.3.1, we easily reduce the proof of Proposition 3.3.5 to that of the following assertion: if $\mathfrak{f} : \mathcal{O}_{\mathcal{E}} \otimes_{\mathfrak{S}_F} \mathfrak{M} \to \mathcal{O}_{\mathcal{E}} \otimes_{\mathfrak{S}_F} \mathfrak{M}'$ is a morphism in $\mathrm{Mod}_{\mathcal{O}_{\mathcal{E}}}^{\varphi}$ then $\mathfrak{f}(\mathfrak{M}) \subset \mathfrak{M}'$. By applying Lemma 3.3.4 to $\mathfrak{f}(\mathfrak{M}) + \mathfrak{M}'$, we may further reduce the proof to that of the following statement: if $\mathfrak{M} \subset \mathfrak{M}' \subset \mathcal{O}_{\mathcal{E}} \otimes_{\mathfrak{S}_F} \mathfrak{M}$ then $\mathfrak{M} = \mathfrak{M}'$. Writing $d := \mathrm{rk}_{\mathfrak{S}_F}(\mathfrak{M}) = \mathrm{rk}_{\mathfrak{S}_F}(\mathfrak{M}')$ and applying $\wedge^d$, we may reduce to the case $d = 1$, and now calculate with bases. Let e (resp. e') be an $\mathfrak{S}_F$-basis of $\mathfrak{M}$ (resp. $\mathfrak{M}'$), and let $a \in \mathfrak{S}_F$ be the unique element with $e = ae'$. Since $\mathcal{O}_{\mathcal{E}} \otimes_{\mathfrak{S}_F} \mathfrak{M} = \mathcal{O}_{\mathcal{E}} \otimes_{\mathfrak{S}_F} \mathfrak{M}'$, by Weierstrass preparation, we may modify our choices of e and e' to assume that $a = A(u) = u^s + c_{s-1}u^{s-1} + \cdots + c_1 u + c_0$ with $c_i \in \varpi \mathcal{O}_{F_0}$. As in Example 3.3.2, we may suppose that $\varphi(e') = \gamma' E^n e'$ and $\varphi(e) = \gamma E^{n'} e$ with $\gamma, \gamma' \in \mathfrak{S}_F^{\times}$. Then

$$\gamma E^{n'} A(u) e' = \gamma E^{n'} e = \varphi(e) = \varphi(A(u))\varphi(e') = \varphi(A(u))\gamma' E^n e'$$

which necessitates $\gamma A(u) E^{n'} = \gamma' \varphi(A(u)) E^n$. Reducing modulo ϖ and comparing u-degrees, we see easily that $n' \geq n$. We therefore have

$$\gamma_0 A(u) E^{n'-n} = \varphi(A(u)) \quad \text{for} \quad \gamma_0 = \gamma(\gamma')^{-1} \in \mathfrak{S}_F^{\times}. \tag{3.3.1}$$

As γ_0 is a unit, it follows from (3.3.1) that $A(u)E^{n'-n}$ and $\varphi(A(u))$ must have the same roots. Since $A(u)$, $\varphi(A(u))$ and E are monic polynomials with roots either 0 or with positive valuation, we conclude that $A(u)E^{n'-n} = \varphi(A(u))$. Let us put $A(u) = u^l A_0(u)$ with $A_0(0) \neq 0$ and $m = n' - n$. Then (3.3.1) simplifies to

$$A_0(u) E^m = (f(u)/u)^l \varphi(A_0(u)). \tag{3.3.2}$$

We first treat the case $l = 0$ (so $A = A_0$); we will then reduce the general case to this one. Put $A^{\varphi}(u) := u^s + \varphi(c_{s-1})u^{s-1} + \cdots + \varphi(c_1)u + \varphi(c_0)$. There is then a bijection between the roots of $A^{\varphi}(u)$ and the roots of $A(u)$ which preserves valuation. Let x_0 be a nonzero root of $A(u)$ which achieves the maximal valuation. Then $A(u)E^m = \varphi(A(u))$ implies that x_0 is root of $\varphi(A(u)) = A^{\varphi}(f(u))$. That is $f(x_0)$ is a root of $A^{\varphi}(u)$. If $f(x_0) \neq 0$, then since $f(u) \equiv u^p \bmod \varpi$ and x_0 has positive valuation, we have $v(f(x_0)) > v(x_0)$, so there exists a root of $A(u)$ with valuation strictly greater than $v(x_0)$, which contradicts our choice of x_0. We must therefore have that $f(x_0) = 0$ is root of $A^{\varphi}(u)$, which contradicts our assumption that $A(0) \neq 0$ ($l = 0$). We conclude that $A(u) = A_0(u)$ has degree zero, and hence that $\mathfrak{M} = \mathfrak{M}'$ as desired.

Now suppose that $l > 0$ and let $r_1 \geq 0$ be the unique integer satisfying $(f(u)/u)^l = E^{r_1} h_1(u)$ for some unique monic $h_1 \in \mathfrak{S}_F$ with $E \nmid h_1(u)$. Comparing u-degrees in (3.3.2) gives $r_1 \leq m$, so $h_1(u) | A_0(u)$ and we have

$A_0(u) = h_1(u)A_1(u)$ for a unique monic polynomial A_1 and $m_1 := m - r_1 \geq 0$. Equation (3.3.2) then becomes

$$A_1(u)E^{m_1} = \varphi(h_1(u))\varphi(A_1(u)).$$

Now let $r_2 \geq 0$ be the unique integer with $\varphi(h_1(u)) = E^{r_2}h_2(u)$ for h_2 a monic polynomial with $E \nmid h_2(u)$, and write $A_1(u) = h_2(u)A_2(u)$ with A_2 monic and $m_2 := m_1 - r_2 \geq 0$. We then have

$$A_2(u)E^{m_2} = \varphi(h_2(u))\varphi(A_2(u)).$$

We continue in this manner, constructing nonnegative integers r_n, m_n with $m_{n+1} := m_n - r_n$ and monic $A_n, h_n \in \mathfrak{S}_F$ with $E \nmid h_n$, $h_nE^{r_n} = \varphi(h_{n-1})$ and $A_{n-1} = h_nA_n$ satisfying the equation

$$A_n(u)E^{m_n} = \varphi(h_n(u))\varphi(A_n(u)). \tag{3.3.3}$$

So long as h_n and A_n are non-constant, we have $\deg A_n < \deg A_{n-1}$, which can not continue indefinitely. We conclude that there is some $n \geq 1$ with either h_n or A_n constant, which forces $h_n = 1$ or $A_n = 1$ by monicity. In the latter case, (3.3.3) implies that $h_{n+1} = 1$, so in any case there is some $n > 0$ with $h_n = 1$. By the construction of the h_m, we then have

$$\varphi^{n-1}((f(u)/u)^l) = \prod_{m=1, r_m \neq 0}^{n} \varphi^{n-m}(E^{r_m}). \tag{3.3.4}$$

We claim that in fact there is only one m with $r_m \neq 0$. Indeed, if there exist $m_1 > m_2$ with $r_{m_i} \neq 0$ for $i = 1, 2$, then writing $f_0(u) = f(u)/u$, we see that $f_0(f^{(m_i)}(\pi)) = 0$ for $i = 1, 2$. Since $f(u) = f_0(u)u$, this implies that $f^{(m_2+1)}(\pi) = 0$. Then

$$0 = f_0(f^{(m_1)}(\pi)) = f_0(f^{(m_1-m_2-1)}(f^{(m_2+1)}(\pi))) = f_0(f^{(m_1-m_2-1)}(0)) = f_0(0),$$

which implies that $u | f_0(u)$. But this contradicts (3.3.2) because $u \nmid A_0(u)$. We conclude that there is a unique m such that $r_m \neq 0$, and it follows from (3.3.4) that there exists $n \geq 0$ such that $\varphi^n(f(u)/u)$ is a power of $E(u)$, contradicting our hypothesis. We must therefore in fact have $l = 0$, whence $\mathfrak{M} = \mathfrak{M}'$ as we showed above. $\square$

Remark 3.3.6. The assumption that $\varphi^{(n)}(f(u)/u)$ is not a power of E for any $n \geq 0$ is satisfied in many cases of interest. For example, it is always satisfied when $a_1 = 0$ (which includes the classical situation $f(u) = u^p$), as then $f(u)/u$ has no constant term while any power of $E = E(u)$ has nonzero constant term.

EXAMPLE 3.3.7. The hypothesis of Proposition 3.3.5 that $\varphi^{(n)}(f(u)/u)$ is not a power of E for any $n \geq 0$ is necessary, as the following examples show:

(1) Fix r, let $0 \leq l \leq r$ be an integer and suppose that we have $\varphi^{(n)}(f(u)/u) = E^l$. Setting $A(u) := f(u) \cdot \varphi(f(u)/u) \cdots \varphi^{n-1}(f(u)/u)$ if $n > 0$ and $A(u) = u$ if $n = 0$, we have $AE^l = \varphi(A)$. In particular, definining $\mathfrak{M} = A(u)\mathfrak{S}_F$ and $\mathfrak{M}' := \mathfrak{S}_F$, we have $\mathfrak{M} \subseteq \mathfrak{M}'$ and both

$\mathfrak{M}$ and $\mathfrak{M}'$ are objects of $\mathrm{Mod}_{\mathfrak{S}_F}^{\varphi,r}$ with height $l \leq r$. However, $\mathfrak{M} \neq \mathfrak{M}'$ and it follows that the conclusion of Proposition 3.3.5 does not hold.

(2) Concretely, recall the situation in Example 2.1.2 where $K = \mathbb{Q}_p(\zeta_p)$, $F = \mathbb{Q}_p$, $\pi = \zeta_p - 1$ and $\varphi(u) = (u+1)^p - 1$. In this case, $E = \varphi(u)/u$, and the Kisin modules $\mathfrak{M}' := \mathfrak{S}_F$ and $\mathfrak{M} := u\mathfrak{S}_F$ are both of height 1 and are non-isomorphic, but $T_{\mathfrak{S}}(\mathfrak{M}) \simeq T_{\mathfrak{S}}(\mathfrak{M}')$.

(3) As a less familiar variant, we can take $\varphi(u) = (u-p)^{p-1}u$ and let $E = \varphi(u) - p$. Then $\varphi(f(u)/u) = E^{p-1}$, and the construction of (1) provides a counterexample.

Corollary 3.3.8. *Suppose that $\varphi^n(f(u)/u)$ is not a power of E for any $n \geq 0$ and $\psi : V' \to V$ is a morphism of height-r representations. Then there are exact sequences*

$$0 \longrightarrow \mathfrak{L} \longrightarrow \mathfrak{M} \longrightarrow \mathfrak{N} \longrightarrow 0, \quad \textit{and} \quad 0 \longrightarrow \mathfrak{N} \longrightarrow \mathfrak{M}' \longrightarrow \mathfrak{N}' \longrightarrow 0$$

in $\mathrm{Mod}_{\mathfrak{S}_F}^{\varphi,r}$ which correspond via $V_{\mathfrak{S}}(\cdot)$ to the exact sequences in $\mathrm{Rep}_F(G_{\underline{\pi}})$:

$$0 \longrightarrow \psi(V') \longrightarrow V \longrightarrow V/\psi(V') \longrightarrow 0,$$

and

$$0 \longrightarrow \ker(\psi) \longrightarrow V' \xrightarrow{\psi} \psi(V') \longrightarrow 0.$$

Proof. We may and do select $G_{\underline{\pi}}$-stable $\mathcal{O}_F$-lattices $T \subseteq V$ and $T' \subseteq V'$ with $\psi(T') \subseteq T$ and $T/\psi(T')$ torsion-free. Thanks to Proposition 3.3.3, there exist $\mathfrak{M}$ and $\mathfrak{M}'$ in $\mathrm{Mod}_{\mathfrak{S}_F}^{\varphi,r}$ with $T = T_{\mathfrak{S}}(\mathfrak{M})$ and $T' = T_{\mathfrak{S}}(\mathfrak{M}')$, and we define $M := \mathcal{O}_{\mathcal{E}} \otimes_{\mathfrak{S}_F} \mathfrak{M}$ and $M' := \mathcal{O}_{\mathcal{E}} \otimes_{\mathfrak{S}_F} \mathfrak{M}'$ and write $\mathfrak{f} : M \to M'$ for the unique morphism in $\mathrm{Mod}_{\mathcal{O}_{\mathcal{E}}}^{\varphi}$ with $\underline{T}(\mathfrak{f}) = \psi|_{T'}$. Let $N' := M'/\mathfrak{f}(M)$ and write $\mathfrak{g} : M' \to N'$ for the natural projection. Writing $N := \mathfrak{f}(M) = \ker \mathfrak{g}$, we then have exact sequences in $\mathrm{Mod}_{\mathcal{O}_{\mathcal{E}}}^{\varphi}$

$$0 \longrightarrow \ker(\mathfrak{f}) \longrightarrow M \xrightarrow{\mathfrak{f}} N \longrightarrow 0$$

and

$$0 \longrightarrow N \longrightarrow M' \xrightarrow{\mathfrak{g}} N' \longrightarrow 0$$

which correspond, via $\underline{T}(\cdot)$, to the exact sequences in $\mathrm{Rep}_{\mathcal{O}_F}(G_{\underline{\pi}})$

$$0 \longrightarrow \psi(T') \longrightarrow T \longrightarrow T/\psi(T') \longrightarrow 0,$$

and

$$0 \longrightarrow \ker(\psi) \longrightarrow T' \xrightarrow{\psi} \psi(T') \longrightarrow 0.$$

Since N' corresponds to $T/\psi(T')$, which is torsion-free, it follows that N' is also torsion free and hence finite and free as an $\mathcal{O}_{\mathcal{E}}$-module. Define $\mathfrak{N} := \ker(\mathfrak{g}) \cap \mathfrak{M}'$, the intersection taken inside of M'. We claim that $\mathfrak{N}$ is an object in $\mathrm{Mod}_{\mathfrak{S}_F}^{\varphi,r}$. First note that by [Fon90, B 1.3.5], the fact that $\mathfrak{M}'$ has height r implies that both $\mathfrak{g}(\mathfrak{M}')$ and $\mathfrak{N}$ have height r, and we need only show that $\mathfrak{N}$ is free over $\mathfrak{S}_F$. To do this, it suffices by Lemma 3.3.4 to prove that $\mathfrak{N} = \mathfrak{N}[1/p] \cap N$ inside

$\mathcal{E}\otimes_{\mathcal{O}_\mathcal{E}} N$, or equivalently that $\mathfrak{N}[1/p]\cap N \subseteq \mathfrak{N}$. For any $x \in \mathfrak{N}[1/p]\cap N$, we have by the very definition of $\mathfrak{N}$ that $x \in \mathfrak{M}'[1/p] \cap M' = \mathfrak{M}'$. As $x \in N = \ker \mathfrak{g}$, we then have $x \in \ker \mathfrak{g} \cap \mathfrak{M}' = \mathfrak{N}'$ as desired. A similar argument shows that $\mathfrak{L} := \ker(\mathfrak{f}) \cap \mathfrak{M}$ is a Kisin module in $\mathrm{Mod}^{\varphi,r}_{\mathfrak{S}_F}$ as well.

Again using Lemma 3.3.4, both $\widetilde{\mathfrak{N}} := \mathfrak{f}(\mathfrak{M})[1/p] \cap N$ and $\mathfrak{N}' := \mathfrak{g}(\mathfrak{M}')[1/p] \cap N'$ are objects of $\mathrm{Mod}^{\varphi,r}_{\mathfrak{S}_F}$. As $\mathcal{O}_\mathcal{E} \otimes_{\mathfrak{S}_F} \widetilde{\mathfrak{N}} = N = \mathcal{O}_\mathcal{E} \otimes_{\mathfrak{S}_F} \mathfrak{N}$, it follows from Proposition 3.3.5 that $\widetilde{\mathfrak{N}} = \mathfrak{N}$. We therefore have exact sequences

$$0 \to \mathfrak{L}[1/p] \to \mathfrak{M}[1/p] \xrightarrow{\mathfrak{f}} \mathfrak{N}[1/p] \to 0,$$

and

$$0 \to \mathfrak{N}[1/p] \to \mathfrak{M}'[1/p] \xrightarrow{\mathfrak{g}} \mathfrak{N}'[1/p] \to 0 \ .$$

Unfortunately, it need not be true in general that $\mathfrak{f}(\mathfrak{M}) = \mathfrak{N}$ or $\mathfrak{g}(\mathfrak{M}') = \mathfrak{N}'$. To remedy this defect, we modify $\mathfrak{M}$ and $\mathfrak{M}'$ as follows. Using the inclusion $\mathfrak{N} \subseteq \mathfrak{M}'$ and the above exact sequences, we may select an $\mathfrak{S}_F[1/p]$-basis $\mathfrak{e}_1, \ldots, \mathfrak{e}_s, \mathfrak{e}_{s+1}, \ldots, \mathfrak{e}_d$ of $\mathfrak{M}[1/p]$ with the property that $\mathfrak{e}_1, \ldots, \mathfrak{e}_s$ is an $\mathfrak{S}_F$-basis of $\mathfrak{L}$ and $\mathfrak{e}'_{s+1} := \mathfrak{f}(\mathfrak{e}_{s+1}), \ldots, \mathfrak{e}'_d := \mathfrak{f}(\mathfrak{e}_d)$ is an $\mathfrak{S}_F$-basis of $\mathfrak{N}$. We may further complete $\mathfrak{e}'_{s+1}, \ldots, \mathfrak{e}'_d$ to an $\mathfrak{S}_F[1/p]$-basis $\mathfrak{e}'_{s+1}, \ldots, \mathfrak{e}'_d, \mathfrak{e}'_{d+1}, \ldots, \mathfrak{e}'_{d'}$ of $\mathfrak{M}'$ with the property that $\mathfrak{e}'_{d+1}, \ldots, \mathfrak{e}'_{d'}$ projects via $\mathfrak{g}$ to an $\mathfrak{S}_F$-basis of $\mathfrak{N}'$. We then have matrix equations

$$\varphi(\mathfrak{e}_1, \ldots, \mathfrak{e}_d) = (\mathfrak{e}_1, \ldots, \mathfrak{e}_d) \begin{pmatrix} A & C \\ 0 & A' \end{pmatrix}$$

and

$$\varphi(\mathfrak{e}'_{s+1}, \ldots, \mathfrak{e}'_d, \mathfrak{e}'_{d+1}, \ldots, \mathfrak{e}'_{d'}) = (\mathfrak{e}'_{s+1}, \ldots, \mathfrak{e}'_d, \mathfrak{e}'_{d+1}, \ldots, \mathfrak{e}'_{d'}) \begin{pmatrix} B & D \\ 0 & B' \end{pmatrix},$$

where the entries of A, A', B, B' are in $\mathfrak{S}_F$, while the entries of C and D are in $\mathfrak{S}_F[1/p]$. Let $m \geq 0$ be such that $p^m C$ and $p^m D$ have all entries in $\mathfrak{S}_F$. Replacing $\mathfrak{M}$ by the $\mathfrak{S}_F$-submodule of $\mathfrak{M}[1/p]$ generated over $\mathfrak{S}_F$ by $p^{-m}\mathfrak{e}_1, \ldots, p^{-m}\mathfrak{e}_s, \mathfrak{e}_{s+1}, \ldots, \mathfrak{e}_d$, and $\mathfrak{M}'$ by the $\mathfrak{S}_F$-submodule of $\mathfrak{M}'[1/p]$ generated by $(\mathfrak{e}'_{s+1}, \ldots, \mathfrak{e}'_d, p^m\mathfrak{e}'_{d+1}, \ldots, p^m\mathfrak{e}'_{d'})$ does the trick. □

4. Constructing Kisin modules from F-crystalline representations

In this section, we associate to any F-crystalline representation a Kisin module in the sense of §3.3 and employ our construction to prove Theorems 1.0.1 and 1.0.2. Throughout, and especially in §4.2, we make free use of many of the ideas of [Kis06] and [KR09]. To surmount the difficulty that we do not in general have a natural N_∇-structure (see the introduction), we will compare our modules over the Robba ring to those of Kisin's classical setting in in §4.4, which will allow us to descend these modules to the desired Kisin modules. The proofs of our main results (Theorems 1.0.1 and 1.0.2) occupies §4.5.

4.1. Generalities on F-crystalline representations. Let V be an F-linear representation of $G = G_K$ or of $G_{\underline{\pi}}$. We write $V^\vee$ for the F-linear dual of V with its natural G or $G_{\underline{\pi}}$-action. We warn the reader at the outset that our notational conventions regarding Fontaine's functors are dual to the standard ones; we have chosen to depart from tradition here as it will be more convenient to deal with the integral theory.

Let V be an object of $\mathrm{Rep}_F(G)$. Then $D_{\mathrm{dR}}(V) := (V^\vee \otimes_{\mathbb{Q}_p} B_{\mathrm{dR}})^G$ is naturally a module over the semilocal ring $K_F := K \otimes_{\mathbb{Q}_p} F$, so we have a decomposition

$$D_{\mathrm{dR}}(V) = \prod_{\mathfrak{m}} D_{\mathrm{dR}}(V)_{\mathfrak{m}}$$

with the product running over all maximal ideals of K_F. We give each $D_{\mathrm{dR}}(V)_{\mathfrak{m}}$ the filtration induced from that of $D_{\mathrm{dR}}(V)$, and we denote by $\mathfrak{m}_0$ the kernel of the natural map $K \otimes_{\mathbb{Q}_p} F \to K$ coming from the given inclusion $F \hookrightarrow K$ and multiplication. Following [KR09], we define:

Definition 4.1.1. We say that $V \in \mathrm{Rep}_F(G)$ is *F-crystalline* if it is crystalline (as a $\mathbb{Q}_p$-linear G-representation) and the filtration on $D_{\mathrm{dR}}(V)_{\mathfrak{m}}$ is trivial ($\mathrm{Fil}^j D_{\mathrm{dR}}(V)_{\mathfrak{m}} = 0$ if $j > 0$ and $\mathrm{Fil}^0 D_{\mathrm{dR}}(V)_{\mathfrak{m}} = D_{\mathrm{dR}}(V)_{\mathfrak{m}}$) when $\mathfrak{m} \neq \mathfrak{m}_0$. We write $\mathrm{Rep}_F^{F\text{-cris}}(G)$ for the category of F-crystalline F-representations of G.

We now wish to describe the category of F-crystalline G-representations in terms of filtered φ-modules. To do this, we define:

Definition 4.1.2. Let $\mathrm{MF}^{\varphi}_{F_0,K}$ be the category of triples $(D, \varphi, \mathrm{Fil}^j D_{F_0,K})$ where D is a finite dimensional F_0-vector space, $\varphi : D \to D$ is a semilinear (over the F-linear extension φ of the p-power Frobenius map $K_0 \to K_0$) endomorphism whose linearization is an F_0-linear isomorphism, and $\mathrm{Fil}^j D_{F_0,K}$ is a separated and exhaustive descending filtration by K-subspaces on the scalar extension $D_{F_0,K} := D \otimes_{F_0} K$. Morphisms in this category are φ-compatible F_0-linear maps $D \to D'$ which are filtration-compatible after applying $\otimes_{F_0} K$.

Let V be an F-crystalline G-representation with F-dimension d. Then

$$D := D_{\mathrm{cris}}(V) := (V^\vee \otimes_{\mathbb{Q}_p} B_{\mathrm{cris}})^G$$

is naturally a module over $F \otimes_{\mathbb{Q}_p} K_0$, equipped with a semilinear (over $1 \otimes \sigma$ for σ the p-power Frobenius automorphism of K_0) Frobenius endomorphism $\varphi : D \to D$ which linearizes to an isomorphism. By our assumption that $K_0 \cap F = \mathbb{Q}_p$, the natural multiplication map $F \otimes_{\mathbb{Q}_p} K_0 \to FK_0 =: F_0$ is an isomorphism, so D is an F_0-vector space which, as V is crystalline as a $\mathbb{Q}_p$-representation, has K_0-dimension $d[F : \mathbb{Q}_p]$, so must have F_0-dimension d. The natural injective map

$$D \otimes_{K_0} K = D_{\mathrm{cris}}(V) \otimes_{K_0} K \hookrightarrow D_{\mathrm{dR}}(V) := (V^\vee \otimes_{\mathbb{Q}_p} B_{\mathrm{dR}})^G$$

is necessarily an isomorphism of $F_K := F \otimes_{\mathbb{Q}_p} K$-modules, so since V is F-crystalline we have a direct sum decomposition of filtered K-vector spaces

$D \otimes_{K_0} K = \bigoplus_{\mathfrak{m}} D_{K,\mathfrak{m}}$, with $D_{K,\mathfrak{m}}$ having trivial filtration unless $\mathfrak{m} = \mathfrak{m}_0$. Noting the canonical identification

$$D_{K,\mathfrak{m}_0} = D \otimes_{F_0} K =: D_{F_0,K},$$

we therefore obtain a filtration on $D_{F_0,K}$. In this way we obtain an object

$$D_{\mathrm{cris},F}(V) := (D, \varphi, \mathrm{Fil}^j D_{F_0,K})$$

of $\mathrm{MF}^{\varphi}_{F_0,K}$.

Conversely, if D is any object of $\mathrm{MF}^{\varphi}_{F_0,K}$, we define

$$V_{\mathrm{cris},F}(D) := \mathrm{Hom}_{F_0,\varphi}(D, B^+_{\mathrm{cris},F}) \cap \mathrm{Hom}_{K,\mathrm{Fil}^\bullet}(D_{F_0,K}, B^+_{\mathrm{cris}} \otimes_{K_0} K)$$

with the intersection taken inside of $\mathrm{Hom}_K(D_{F_0,K}, B^+_{\mathrm{cris}} \otimes_{K_0} K)$, via the map

$$\mathrm{Hom}_{F_0,\varphi}(D, B^+_{\mathrm{cris},F}) \hookrightarrow \mathrm{Hom}_K(D_{F_0,K}, B^+_{\mathrm{cris}} \otimes_{K_0} K)$$

that sends an F_0-linear $h : D \to B^+_{\mathrm{cris},F}$ to its linear extension along $F_0 \to K$.

Proposition 4.1.3. *Let $V \in \mathrm{Rep}^{F\text{-cris}}_F(G)$. Then $V_{\mathrm{cris},F}(D_{\mathrm{cris},F}(V)) \simeq V$.*

Proof. Set $D = D_{\mathrm{cris}}(V) := (V^\vee \otimes_{\mathbb{Q}_p} B^+_{\mathrm{cris}})^G$ and put $D_K := K \otimes_{K_0} D$. As it is well-known (e.g. [Fon94b, §5.3.7]) that $V \simeq V_{\mathrm{cris}}(D)$ as $F[G]$-modules, for

$$V_{\mathrm{cris}}(D) = \mathrm{Hom}_{K_0,\varphi}(D, B^+_{\mathrm{cris}}) \cap \mathrm{Hom}_{K,\mathrm{Fil}}(D_K, B^+_{\mathrm{cris}} \otimes_{K_0} K),$$

it is enough to prove that $V_{\mathrm{cris}}(D) \simeq V_{\mathrm{cris},F}(D)$ as $F[G]$-modules. We will first construct an F-linear isomorphism

$$\iota : \mathrm{Hom}_{K_0}(D, B^+_{\mathrm{cris}}) \xrightarrow{\simeq} \mathrm{Hom}_{F_0}(D, B^+_{\mathrm{cris},F}). \tag{4.1.1}$$

Writing $D_{F_0} = D \otimes_{K_0} F_0$, which is an $F_0 \otimes_{K_0} F_0$-module, we note that $F_0 \simeq F_0 \otimes_{F_0} F_0$ is a subfield of $F_0 \otimes_{K_0} F_0$, so we may and do regard $D \simeq D \otimes_{F_0} F_0$ as an F_0-subspace of D_{F_0}. Thus, restricting homomorphisms from D_{F_0} to the subspace D gives a natural map of F-vector spaces $\iota' : \mathrm{Hom}_{F_0}(D_{F_0}, B^+_{\mathrm{cris},F}) \to \mathrm{Hom}_{F_0}(D, B^+_{\mathrm{cris},F})$. As $\mathrm{Hom}_{K_0}(D, B^+_{\mathrm{cris}})$ is easily checked to be an F-subspace of $\mathrm{Hom}_{F_0}(D_{F_0}, B^+_{\mathrm{cris},F})$, restriction of ι' to $\mathrm{Hom}_{K_0}(D, B^+_{\mathrm{cris}})$ then gives the desired map (4.1.1).

To check that (4.1.1) is an isomorphism, we explicitly compute with bases: Let $e_1, \ldots, e_d$ be an F_0-basis of D and $\beta_1, \ldots, \beta_{e_F}$ a K_0-basis of F_0. Any $x \in D$ can then be uniquely expressed as a linear combination $x = \sum_{ij} a_{ij}\beta_j e_i$ for some $a_{ij} \in K_0$, while any $y \in D_{F_0}$ admits a unique representation of the form $y = \sum_{i,j,l} a_{ijl}\beta_j e_i \otimes \beta_l$ with $a_{ijl} \in K_0$. The natural F-linear inclusion $D \hookrightarrow D_{F_0}$ induced by $F_0 \otimes_{F_0} F_0 \subset F_0 \otimes_{K_0} F_0$ carries $x \in D$ above to

$$x = \sum_{ij} a_{ij}\beta_j e_i \otimes \beta_j \in D_{F_0}.$$

In particular, if $h \in \mathrm{Hom}_{K_0}(D, B^+_{\mathrm{cris}})$, then h is uniquely determined by the matrix $\{c_{ij}\}$ with $c_{ij} := h(\beta_j e_i) \in B^+_{\mathrm{cris}}$, and it follows from definitions that

$\iota(f)(x) = \sum_{ij} a_{ij}c_{ij} \otimes \beta_j$ as an element of $B^+_{\mathrm{cris}} \otimes_{K_0} F_0$. From this explicit description of ι, one checks easily that ι is indeed an isomorphism.

From the very definition of (4.1.1), one checks that ι induces an isomorphism

$$\mathrm{Hom}_{K_0,\varphi}(D, B^+_{\mathrm{cris}}) \simeq \mathrm{Hom}_{F_0,\varphi}(D, B^+_{\mathrm{cris},F}),$$

so to complete the proof it remains to show that for any $h \in \mathrm{Hom}_{K_0}(D, B^+_{\mathrm{cris}})$, the scalar extension $h\otimes 1 : D\otimes_{K_0} K \to B^+_{\mathrm{cris}}\otimes_{K_0} K$ is compatible with filtrations if and only if this is true of $\iota(h) \otimes 1 : D \otimes_{F_0} K \to B^+_{\mathrm{cris},F} \otimes_{F_0} K$. Observe that the construction of the map (4.1.1) gives the following commutative diagram,

$$\begin{tikzcd} D_K \arrow[r, "h\otimes_{K_0}1"] & B^+_{\mathrm{cris}} \otimes_{K_0} K \\ D_{F_0,K} \arrow[u, hook] \arrow[r, "\iota(h)\otimes_{F_0}1"] & B^+_{\mathrm{cris},F} \otimes_{F_0} K \arrow[u, equal] \end{tikzcd} \tag{4.1.2}$$

where we make the identification $B^+_{\mathrm{cris}} \otimes_{K_0} K = B^+_{\mathrm{cris},F} \otimes_{F_0} K$. As V is F-crystalline, we have $\mathrm{Fil}^i D_K = \mathrm{Fil}^i D_{F_0,K}$ for $i \geq 1$ by definition, and it follows from this and (4.1.2) that

$$(h \underset{K_0}{\otimes} 1)(\mathrm{Fil}^i D_K) \subset \mathrm{Fil}^i B^+_{\mathrm{cris}} \underset{K_0}{\otimes} K \iff (\iota(h) \underset{F_0}{\otimes} 1)(\mathrm{Fil}^i D_{F_0,K}) \subset \mathrm{Fil}^i B^+_{\mathrm{cris},F} \underset{F_0}{\otimes} K,$$

which completes that proof of $V_{\mathrm{cris},F}(D) \simeq V_{\mathrm{cris}}(D) \simeq V$ as $F[G]$-modules. $\square$

Let V be an F-linear representation of G. For each field embedding $\tau : F \to \overline{K}$, we define the set *$\tau$-Hodge-Tate weights* of V:

$$\mathrm{HT}_\tau(V) := \{i \in \mathbb{Z} \mid (V \otimes_{F,\tau} \mathbb{C}_K(-i))^G \neq \{0\}\},$$

where $\mathbb{C}_K$ is the p-adic completion of $\overline{K}$. It is easy to see that V is F-crystalline if and only if V is crystalline and $\mathrm{HT}_\tau(V) = \{0\}$ unless τ is the trivial embedding $\tau_0 : F \subset K \subset \overline{K}$. For the remainder of this paper, we will *fix* a nonnegative integer r with the property that $\mathrm{HT}_{\tau_0}(V) \subset \{0, \ldots, r\}$, or equivalently, $\mathrm{Fil}^{r+1} D_{F_0,K} = \{0\}$. We denote by $\mathrm{Rep}_F^{F\text{-cris},r}(G)$ the category of F-crystalline representations V of G with $\mathrm{HT}_{\tau_0}(V) \subset \{0, \ldots, r\}$.

4.2. φ-modules over $\mathfrak{O}$. Recall that we equip $F_0(\!(u)\!)$ with the Frobenius endomorphism $\varphi : F_0(\!(u)\!) \to F_0(\!(u)\!)$ which acts as the canonical Frobenius on K_0, acts trivially on F, and sends u to $f(u)$. For any sub-interval $I \subset [0, 1)$, we write $\mathfrak{O}_I$ for the subring of $F_0(\!(u)\!)$ consisting of those elements which converge for all $x \in \overline{K}$ with $|x| \in I$. For ease of notation, we put $\mathfrak{O} = \mathfrak{O}_{[0,1)}$ and as before we set $K_n = K(\pi_{n-1})$. We denote by $\widehat{\mathfrak{S}}_n$ the completion of $K_{n+1}\otimes_{F_0}\mathfrak{S}_F$ at the maximal ideal $(u-\pi_n)$. Equip $\widehat{\mathfrak{S}}_n$ with its $(u-\pi_n)$-adic filtration, which extends to a filtration on the quotient field $\mathrm{Fr}\,\widehat{\mathfrak{S}}_n = \widehat{\mathfrak{S}}_n[1/(u - \pi_n)]$. Note that for any n we have natural maps of F_0-algebras $\mathfrak{S}_F[1/p] \hookrightarrow \mathfrak{O} \hookrightarrow \widehat{\mathfrak{S}}_n$ given by sending u to u, where the first map has dense image. We will write $\varphi_W : \mathfrak{S}_F \to \mathfrak{S}_F$ for the $\mathcal{O}_F[\![u]\!]$-linear map which acts on $W(k)$ via the canonical lift of Frobenius, and by $\varphi_{\mathfrak{S}/W} : \mathfrak{S}_F \to \mathfrak{S}_F$ the $\mathcal{O}_{F_0}$-linear map which sends u to $f(u)$. Let $c_0 = E(0) \in F_0$ and set

$$\lambda := \prod_{n=0}^{\infty} \varphi^n(E(u)/c_0) \in \mathfrak{O}.$$

A *φ-module over $\mathfrak{O}$* is a finite free $\mathfrak{O}$-module $\mathcal{M}$ equipped with a semilinear endomorphism $\varphi_{\mathcal{M}} : \mathcal{M} \to \mathcal{M}$. We say that $\mathcal{M}$ is of *finite E-height* if the cokernel of the $\mathfrak{O}$-linear map $1 \otimes \varphi_{\mathcal{M}} : \varphi^*\mathcal{M} \to \mathcal{M}$ is killed by E^r for some r, and we write $\mathrm{Mod}_{\mathfrak{O}}^{\varphi,r}$ for the category of φ-modules over $\mathfrak{O}$ of E-height r. Scalar extension along the inclusion $\mathfrak{S}_F \hookrightarrow \mathfrak{O}$ gives a functor $\mathrm{Mod}_{\mathfrak{S}_F}^{\varphi,r} \to \mathrm{Mod}_{\mathfrak{O}}^{\varphi,r}$ from height-r Kisin modules to φ-modules over $\mathfrak{O}$ of E-height r.

Now let $V \in \mathrm{Rep}_F^{F\text{-cris},r}(G)$ be any F-crystalline representation of G, and write $D := D_{\mathrm{cris},F}(V)$ for the corresponding filtered φ-module. We functorially associate to D an $\mathfrak{O}$-module $\mathcal{M}(D)$ as follows: For each nonnegative integer n, let ι_n be the composite map:

$$\mathfrak{O} \otimes_{F_0} D \xrightarrow{\varphi_W^{-n} \otimes \varphi_D^{-n}} \mathfrak{O} \otimes_{F_0} D \longrightarrow \widehat{\mathfrak{S}}_n \otimes_{F_0} D = \widehat{\mathfrak{S}}_n \otimes_K D_{F_0,K}, \tag{4.2.1}$$

where the second morphism is induced by the canonical inclusion $\mathfrak{O} \to \widehat{\mathfrak{S}}_n$. We again write ι_n for the canonical extension

$$\iota_n : \mathfrak{O}[1/\lambda] \otimes_{F_0} D \longrightarrow \mathrm{Fr}\,\widehat{\mathfrak{S}}_n \otimes_K D_{F_0,K},$$

and we define

$$\mathcal{M}(D) := \left\{ x \in \mathfrak{O}[1/\lambda] \underset{F_0}{\otimes} D \;\middle|\; \iota_n(x) \in \mathrm{Fil}^0 \left(\mathrm{Fr}\,\widehat{\mathfrak{S}}_n \underset{K}{\otimes} D_{F_0,K} \right),\ \forall n \geq 0 \right\}.$$

Proposition 4.2.1. *$\mathcal{M}(D)$ is a φ-module over $\mathfrak{O}$ of E-height r.*

Proof. This is Lemma 1.2.2 in [Kis06] (also see Lemma (2.2.1) in [KR09]) with the following minor modifications: first note that we only discuss crystalline representation here, so we do not need the "logarithm element" ℓ_n which occurs in Kisin's classical setting (strictly speaking, we do not know how to construct ℓ_n in our general setting). Likewise, we may replace $\mathcal{D}_0 := (\mathfrak{O}[\ell_n] \otimes_{K_0} D)^{N=0}$ in the proof of [Kis06, 1.2.2] with $\mathcal{D}_0 = \mathfrak{O} \otimes_{F_0} D$ throughout. In the classical setting, Kisin showed that $\mathcal{M}(D)$ also has an N_∇-structure, which we entirely ignore here (once again, we do not know how to construct N_∇ in general). This is of no harm, as the proof of Lemma 1.2.2 does not use the N_∇-structure of $\mathfrak{O}$ in any way. Finally, we note that Lemma 1.1.4 of [Kis06], which plays an important role in the proof of [Kis06, 1.2.2], is well-known for $\mathfrak{O}$-modules in our more general context.[4] □

[4] Indeed, Kisin's proof of [Kis06, 1.1.4] relies on §4 of Berger's paper [Ber02] as well as results of Lazard [Laz62, §7–8] and Lemma 2.4.1 of [Ked04], while the required facts in [Ber02] build on Lazard's work in a natural way. But [Laz62] already deals in the generality we need, as does Kedlaya [Ked04]. Thus, one checks that all the proofs of the results needed to establish [Kis06, 1.1.4] (as well as Kisin's argument itself) carry over *mutatis mutandis* to our more general situation.

As above, let us write $\mathcal{D}_0 := \mathfrak{O} \otimes_{F_0} D$. We record here the following useful facts, which arise out of (our adaptation of) Kisin's proof of Proposition 4.2.1:

(1) $\mathcal{D}_0 \subset \mathcal{M} \subset \lambda^{-r}\mathcal{D}_0$.

(2) ι_0 induces an isomorphism of $\widehat{\mathfrak{S}}_0$-modules

$$\widehat{\mathfrak{S}}_0 \underset{\mathfrak{O}}{\otimes} \mathcal{M}(D) \simeq \sum_{j\geq 0}(u-\pi)^{-j}\widehat{\mathfrak{S}}_0 \underset{K}{\otimes} \mathrm{Fil}^j D_{F_0,K} = \sum_{j\geq 0} E^{-j}\widehat{\mathfrak{S}}_0 \underset{K}{\otimes} \mathrm{Fil}^j D_{F_0,K}.$$

Consider now the obvious inclusions $D \hookrightarrow \mathcal{D}_0 \subset \mathcal{M}(D)$. As Frobenius induces a linear isomorphism $\varphi^* D \simeq D$, we obtain a linear isomorphism $\varphi^*\mathcal{D}_0 \simeq \mathcal{D}_0$ and hence an injection $\xi : \mathcal{D}_0 \to \varphi^*(\mathcal{M}(D))$. Defining $\mathfrak{O}_\alpha := \mathfrak{S}_F[\![\frac{E^p}{\varpi}]\!][1/p]$, one checks that $\mathfrak{O}_\alpha = \mathfrak{S}_F[\![\frac{u^{e_0 p}}{\varpi}]\!][1/p]$ and that $\mathfrak{O} \subset \mathfrak{O}_\alpha \subset \mathfrak{O}_{[0,|\pi|^{1/p})}$.

Lemma 4.2.2. *The map $\xi_\alpha := \mathfrak{O}_\alpha \otimes_{\mathfrak{O}} \xi : \mathfrak{O}_\alpha \otimes_{F_0} D \to \mathfrak{O}_\alpha \otimes_{\mathfrak{O}} \varphi^*\mathcal{M}(D)$ is an isomorphism.*

Proof. Using the containments $\mathcal{D}_0 \subset \mathcal{M}(D) \subset \lambda^{-r}\mathcal{D}_0$ and $\mathcal{D}_0 \simeq \varphi^*\mathcal{D}_0$, we see that $\mathcal{D}_0 \subset \varphi^*(\mathcal{M}(D)) \subset \varphi(\lambda)^{-r}\mathcal{D}_0$. It is easy to check that $\varphi(\lambda)$ is a unit in $\mathfrak{O}_\alpha$, and it follows that ξ_α is an isomorphism. ☐

For an object $\mathcal{M} \in \mathrm{Mod}_{\mathfrak{O}}^{\varphi,r}$, we define a decreasing filtration on $\varphi^*\mathcal{M}$ by:

$$\mathrm{Fil}^i(\varphi^*\mathcal{M}) := \{x \in \varphi^*\mathcal{M} \mid (1\otimes\varphi)(x) \in E^i\mathcal{M}\}. \tag{4.2.2}$$

On the other hand, using the evident inclusions $\mathfrak{O}_\alpha \subset \mathfrak{O}_{[0,|\pi|^{1/p})} \subset \widehat{\mathfrak{S}}_0$ we obtain a canonical injection $\mathfrak{O}_\alpha \otimes_{F_0} D \hookrightarrow \widehat{\mathfrak{S}}_0 \otimes_K D_{F_0,K}$, which allows us to equip $\mathfrak{O}_\alpha \otimes_{\mathfrak{O}} \mathcal{D}_0$ with the natural subspace filtration, using the tensor product filtration on $\widehat{\mathfrak{S}}_0 \otimes_K D_{F_0,K}$.

Lemma 4.2.3. *The inverse isomorphism*

$$\xi'_\alpha : \mathfrak{O}_\alpha \otimes_{\mathfrak{O}} \varphi^*\mathcal{M}(D) \xrightarrow[(\xi_\alpha)^{-1}]{\simeq} \mathfrak{O}_\alpha \otimes_{F_0} D$$

of Lemma 4.2.2 *is compatible with filtrations and Frobenius.*

Proof. Clearly, ξ'_α is compatible with Frobenius. To prove that ξ'_α is filtration compatible, we use the two facts recorded after Proposition 4.2.1. As noted above, $\varphi(\lambda)$ is a unit in $\widehat{\mathfrak{S}}_0$, so the first fact implies that the injective map $\xi : \mathcal{D}_0 \simeq \varphi^*\mathcal{D}_0 \hookrightarrow \varphi^*\mathcal{M}(D)$ is an isomorphism after tensoring with $\widehat{\mathfrak{S}}_0$. Put $\widehat{\mathcal{D}}_0 := \widehat{\mathfrak{S}}_0 \otimes_{\mathfrak{O}} \mathcal{D}_0$ and define an auxillary filtration on $\widehat{\mathcal{D}}_0$ by

$$\widetilde{\mathrm{Fil}}^i\widehat{\mathcal{D}}_0 := \widehat{\mathcal{D}}_0 \cap E^i(\widehat{\mathfrak{S}}_0 \otimes_{\mathfrak{O}} \mathcal{M}(D)).$$

From the very definition (4.2.2), it is clear that $1\otimes\xi : \widehat{\mathcal{D}}_0 \simeq \widehat{\mathfrak{S}}_0 \otimes_{\mathfrak{O}} \varphi^*\mathcal{M}(D)$ carries $\widetilde{\mathrm{Fil}}^i\widehat{\mathcal{D}}_0$ isomorphically onto $\mathrm{Fil}^i(\widehat{\mathfrak{S}}_0 \otimes_{\mathfrak{O}} \varphi^*\mathcal{M}(D))$. On the other hand, the second fact above implies that an element $d \in \widehat{\mathcal{D}}_0$ lies in $E^i(\widehat{\mathfrak{S}}_0 \otimes_{\mathfrak{O}} \mathcal{M}(D))$ if and only if $\iota_0(d) \in \mathrm{Fil}^i(\widehat{\mathfrak{S}}_0 \otimes_K D_{F_0,K})$, from which $\widetilde{\mathrm{Fil}}^i\widehat{\mathcal{D}}_0 = \mathrm{Fil}^i\widehat{\mathcal{D}}_0$ follows. Hence ξ'_α is indeed compatible with filtrations. ☐

For simplicity, let us put $\mathcal{M} := \mathcal{M}(D)$. It follows from Lemma 4.2.3 that the isomorphism ξ'_α specializes to give a natural identification of φ-modules $D \simeq \varphi^*\mathcal{M}/u\varphi^*\mathcal{M}$ as well as a natural identification of filtered K-vector spaces $D_{F_0,K} \simeq \varphi^*\mathcal{M}/E\varphi^*\mathcal{M}$. Writing ψ_π for the composite mapping

$$\psi_\pi : \varphi^*\mathcal{M} \twoheadrightarrow \varphi^*\mathcal{M}/E\varphi^*\mathcal{M} \simeq D_{F_0,K},$$

we therefore obtain:

Corollary 4.2.4. *The map $\psi_\pi : \varphi^*\mathcal{M}(D) \to D_{F_0,K}$ is filtration compatible.*

Remark 4.2.5. In the classical situation where $F = \mathbb{Q}_p$ and $f(u) = u^p$, to *any* object $\mathcal{M}$ of $\mathrm{Mod}_{\mathfrak{O}}^{\varphi,r}$, Kisin functorially associates a filtered φ-module via $D(\mathcal{M}) := \varphi^*\mathcal{M}/u\varphi^*\mathcal{M}$ with $\mathrm{Fil}^i(D(\mathcal{M})_K) := \psi_\pi(\mathrm{Fil}^i\varphi^*\mathcal{M})$. That this is possible rests crucially on the existence of a unique φ-equivariant isomorphism

$$\xi_\alpha : \mathfrak{O}_\alpha \otimes_{F_0} D(\mathcal{M}) \simeq \mathfrak{O}_\alpha \otimes_{\mathfrak{O}} \varphi^*\mathcal{M}$$

reducing modulo u to the given identification $D(\mathcal{M}) = \varphi^*\mathcal{M}/u\varphi^*\mathcal{M}$, which is Lemma 1.2.6 of [Kis06]. For more general F and $f(u) = u^p + \cdots + a_1u$, we are only able to construct such a map ξ_α under the restriction $\varpi^{r+1}|a_1$ in $\mathcal{O}_F$; see Lemma 4.5.6.

To conclude this section, we record the following further consequence of Lemma 4.2.3: Setting $\widetilde{B}_\alpha := W(R)[\![\frac{E^p}{\varpi}]\!][1/p] \subset B^+_{\mathrm{cris},F}$, one checks that the subspace filtration $\{\mathrm{Fil}^n\widetilde{B}_\alpha\}_n$ coincides with the filtration $\{E^n\widetilde{B}_\alpha\}_n$. As $\mathfrak{S}_F \subset W(R)_F$, we have a canonical inclusion $\mathfrak{O}_\alpha \subset \widetilde{B}_\alpha$, and the map ξ_α of Lemma 4.2.2 induces a natural isomorphism

$$\xi'_{\widetilde{B}_\alpha} : \widetilde{B}_\alpha \otimes_{\mathfrak{O}} \varphi^*\mathcal{M} \simeq \widetilde{B}_\alpha \otimes_{F_0} D. \tag{4.2.3}$$

As the inclusion $\widehat{\mathfrak{S}}_0 \subset B^+_{\mathrm{dR}}$ is compatible with filtrations, we deduce:

Corollary 4.2.6. *The map* (4.2.3) *is compatible with Frobenius and filtrations.*

4.3. The classical setting. For future reference, we now recall the main results in Kisin's classical situation, where $F = \mathbb{Q}_p$ and $f(u) = u^p$. In this subsection only, we fix a choice $\underline{\pi} := \{\pi_n\}_n$ of p-power compatible roots of a fixed uniformizer $\pi = \pi_0$ in K, and set $K_\infty := K_{\underline{\pi}}$ and $G_\infty := G_{\underline{\pi}}$. The following summarizes the main results in this setting:

Theorem 4.3.1 ([Kis06]). *Let V be a $\mathbb{Q}_p$-valued crystalline representation of G with Hodge-Tate weights in $\{0, \ldots, r\}$ and $T \subset V$ a G_∞-stable $\mathbb{Z}_p$-lattice. Then:*

(1) *There exists a unique Kisin module $\mathfrak{M}$ so that $T_{\mathfrak{S}}(\mathfrak{M}) \simeq T$ as $\mathbb{Z}_p[G_\infty]$-modules.*
(2) *If $D = D_{\mathrm{cris}}(V)$ is the associated filtered φ-module, then one has $\mathcal{M}(D) \simeq \mathfrak{O} \otimes_{\mathfrak{S}} \mathfrak{M}$ as φ-modules.*

Proof. These are the main results of [Kis06] restricted to crystalline representations. ◻

Now let F be an arbitrary extension of $\mathbb{Q}_p$ contained in K and let V be an F-crystalline representation and T a G_∞-stable $\mathcal{O}_F$-lattice in V. Viewing V as a crystalline $\mathbb{Q}_p$-valued representation and T as a G_∞-stable $\mathbb{Z}_p$-module, by Theorem 4.3.1 there is a unique (classical) Kisin module $\mathfrak{M}$ attached to T, which is of finite $\widetilde{E}$-height, for $\widetilde{E} := \widetilde{E}(u)$ the minimal polynomial of π over K_0 (we write $\widetilde{E}$ to distinguish this polynomial from our fixed E, which by definition is the minimal polynomial of π over $F_0 = FK_0$). The additional $\mathcal{O}_F$-structure on T is reflected on the classical Kisin module in the following way:

Corollary 4.3.2. *The classical Kisin module $\mathfrak{M}$ is naturally a finite and free $\mathfrak{S}_F$-module and as such has E-height r.*

Proof. Proposition 3.4 of [GLS14] shows that $\mathfrak{M}$ is naturally a finite and free $\mathfrak{S}_F$-module (see also the proof of [Kis08, Prop. 1.6.4]). Factor $\widetilde{E}$ in $\mathcal{O}_{F_0}[u]$ as $\widetilde{E} = E_1 \cdots E_{e_F}$ with $E_1 = E$, and for each i write $\widehat{\mathfrak{S}}_{E_i}$ for the completion of the localization of $\mathfrak{S}_F$ at the ideal (E_i). We must prove that the injective map $1 \otimes \varphi : \varphi^*\mathfrak{M} \to \mathfrak{M}$ has cokernel killed by a power of $E = E_1$. To do this, it suffices to prove that the scalar extension

$$\varphi_i^\sharp : \widehat{\mathfrak{S}}_{E_i} \otimes_{\mathfrak{S}_F} \varphi^*\mathfrak{M} \xrightarrow{1\otimes(1\otimes\varphi)} \widehat{\mathfrak{S}}_{E_i} \otimes_{\mathfrak{S}_F} \mathfrak{M} \tag{4.3.1}$$

of $1\otimes\varphi$ along $\mathfrak{S}_F \to \widehat{\mathfrak{S}}_{E_i}$ is an isomorphism when $i > 1$. Writing $\mathcal{M} := \mathcal{M}(D)$, we recall that the map $\psi_\pi : \varphi^*\mathcal{M} \to D_K$ is compatible with filtrations thanks to Corollary 4.2.4, from which it follows that the map

$$\psi_\pi\big|_{\varphi^*\mathfrak{M}} : \varphi^*\mathfrak{M} \longrightarrow \varphi^*\mathfrak{M}/\widetilde{E}\varphi^*\mathfrak{M} \hookrightarrow D_K$$

is also filtration-compatible. As V is F-crystalline, for any $i > 1$ we have $\mathrm{Fil}^j D_{K,\mathfrak{m}_i} = 0$ for all $j \geq 1$, where $\mathfrak{m}_i$ is the maximal ideal of $F \otimes_{\mathbb{Q}_p} K$ corresponding to E_i, and it follows that $\mathrm{Fil}^1\varphi^*\mathfrak{M} \subset E_i\varphi^*\mathfrak{M}$ for all $i > 1$. We then claim that for $i > 1$ the map $\overline{1\otimes\varphi} : \varphi^*\mathfrak{M}/E_i\varphi^*\mathfrak{M} \to \mathfrak{M}/E_i\mathfrak{M}$ induced from $1 \otimes \varphi$ by reduction modulo E_i is *injective*. To see this, observe that if $x \in \varphi^*\mathfrak{M}$ has $(1 \otimes \varphi)(x) = E_i m$ for $m \in \mathfrak{M}$, then writing $y := \prod_{j\neq i} E_j x$, we have $(1 \otimes \varphi)(y) = \widetilde{E}m$ so that $y \in \mathrm{Fil}^1\varphi^*\mathfrak{M}$ by the very definition of the filtration on $\varphi^*\mathfrak{M}$. By what we have seen above, we then have $y \in E_i\varphi^*\mathfrak{M}$, so since E_i is coprime to $\prod_{j\neq i} E_j$, we obtain $x \in E_i\varphi^*\mathfrak{M}$ as claimed. Now both $\varphi^*\mathfrak{M}$ and $\mathfrak{M}$ are $\mathfrak{S}_F$-free of the same rank, so as $\overline{1\otimes\varphi}$ is injective, we see that $\mathbb{Q}_p \otimes_{\mathbb{Z}_p} \overline{1\otimes\varphi}$ is an isomorphism for $i > 1$. But this map coincides with the map $\overline{\varphi}_i^\sharp$ obtained from (4.3.1) by reduction modulo E_i, so it follows that $\varphi_i^\sharp$ is an isomorphism as well, as desired. $\square$

4.4. Comparing constructions. Let us first recall some standard facts about the Robba ring as in [Kis06]. For finer details of the Robba ring $\mathcal{R}$ and its subring $\mathcal{R}^b$, we refer to §2 (in particular §2.3) of [Ked04], noting that several different notations are commonly used (in particular, we advise the reader that $\mathcal{R}^b = \mathcal{E}^\dagger = \Gamma_{\mathrm{con}}^{k((t))}[1/p]$). The *Robba ring* is defined as

$$\mathcal{R} := \lim_{s\to 1^-} \mathfrak{O}_{(s,1)}$$

and comes equipped with a Frobenius endomorphism, which is induced by the canonical maps $\varphi : \mathfrak{O}_{(s,1)} \to \mathfrak{O}_{(s^{1/p},1)}$. Writing $\mathfrak{O}^b_{(s,1)} \subset \mathfrak{O}_{(s,1)}$ for the subring of functions which are bounded, we also define the *bounded Robba ring*:

$$\mathcal{R}^b := \lim_{s\to 1^-} \mathfrak{O}^b_{(s,1)},$$

which is naturally a Frobenius-stable subring of $\mathcal{R}$. Finally, we put

$$\mathcal{O}_{\mathcal{R}^b} := \left\{ \sum_{n\in\mathbb{Z}} a_n u^n \in \mathcal{R}^b \;\middle|\; a_n \in \mathcal{O}_{F_0}\ \forall n \in \mathbb{Z} \right\};$$

this is a Henselian discrete valuation ring with uniformizer ϖ and residue field $k(\!(u)\!)$. One checks that the fraction field of $\mathcal{O}_{\mathcal{R}^b}$ is $\mathcal{R}^b$, justifying our notation; in particular, $\mathcal{R}^b$ is a field. Note that $\mathfrak{O}$ is canonically a Frobenius-stable subring of $\mathcal{R}$.

By definition, *a φ-module over $\mathcal{R}$* is a finite free $\mathcal{R}$-module $\mathcal{M}$ equipped with a φ-semilinear map $\varphi_{\mathcal{M}} : \mathcal{M} \to \mathcal{M}$ whose linearization $1 \otimes \varphi : \varphi^*\mathcal{M} \to \mathcal{M}$ is an isomorphism. One checks that $E(u) \in \mathfrak{O}$ is a unit in $\mathcal{R}$, so that scalar extension along $\mathfrak{O} \hookrightarrow \mathcal{R}$ gives a functor from φ-modules over $\mathfrak{O}$ to φ-modules over $\mathcal{R}$. A φ-module $\mathcal{M}$ over $\mathcal{R}$ is *étale* if $\mathcal{M}$ admits a basis with the property that the corresponding matrix of $\varphi_{\mathcal{M}}$ lies in $\mathrm{GL}_d(\mathcal{O}_{\mathcal{R}^b})$; by a slight abuse of terminology, we will say that a φ-module over $\mathfrak{O}$ is étale if its scalar extension to $\mathcal{R}$ is. Our main result of this subsection is the following:

Theorem 4.4.1. *Let $V \in \mathrm{Rep}_F^{F\text{-cris},r}(G)$ and write $D := D_{\mathrm{cris},F}(V)$ for the corresponding filtered φ-module. If $\mathcal{M}(D)$ is the φ-module over $\mathfrak{O}$ attached to D as in* §4.2*, we have:*

1. *$\mathcal{M}(D)$ is étale;*
2. *There exists a Kisin module $\mathfrak{M} \in \mathrm{Mod}_{\mathfrak{S}_F}^{\varphi,r}$ such that $\mathfrak{O}\otimes_{\mathfrak{S}_F}\mathfrak{M} \simeq \mathcal{M}(D)$.*

First note that there is a canonical inclusion $\mathfrak{S}_F \hookrightarrow \mathcal{O}_{\mathcal{R}^b}$, so that (2) implies (1). It follows that the above theorem is true in the classical setting of Kisin by Theorem 4.3.1. In what follows, we will reduce the general case of Theorem 4.4.1 to the known instance of it in the classical setting. To ease notation, we will adorn various objects with a superscript of "c" to signify that they are objects in the classical setting. We likewise abbreviate $\mathcal{M} := \mathcal{M}(D)$ and $\mathcal{M}^c := \mathcal{M}^c(D)$. We note that $\mathfrak{O}^c_\alpha \subset \widetilde{B}_\alpha$, as $E(u^c)$ is another generator of $\mathrm{Fil}^1 W(R)_F$ so $E(u^c) = \mu E(u)$ for some $\mu \in W(R)_F^\times$ thanks to Lemma 2.1.3.

By Corollary 4.2.6, the $\widetilde{B}_\alpha$-linear isomorphism $\xi'_{\widetilde{B}_\alpha} : \widetilde{B}_\alpha \otimes_{\mathfrak{O}} \varphi^*\mathcal{M} \simeq \widetilde{B}_\alpha \otimes_{F_0} D$ is Frobenius and filtration compatible. The key point is that the Frobenius and filtration on $\widetilde{B}_\alpha \otimes_{F_0} D$ are *canonical* (recall that the filtration on $\widetilde{B}_\alpha \otimes_{F_0} D$ is induced by the inclusion $\widetilde{B}_\alpha \otimes_{F_0} D \hookrightarrow B^+_{\mathrm{dR}} \otimes_K D_{F_0,K}$) and are independent of the choice of $\varphi(u) = f(u)$. We therefore have a natural isomorphism

$$\widetilde{\xi}: \quad \widetilde{B}_\alpha \otimes_{\mathfrak{O}} \varphi^*\mathcal{M} \simeq \widetilde{B}_\alpha \otimes_{\mathfrak{O}^c} \varphi^*\mathcal{M}^c, \tag{4.4.1}$$

that is Frobenius and filtration compatible.

Lemma 4.4.2. *There is a $\widetilde{B}_\alpha$-linear and Frobenius-compatible isomorphism*

$$\eta : \widetilde{B}_\alpha \otimes_{\mathfrak{O}} \mathcal{M} \simeq \widetilde{B}_\alpha \otimes_{\mathfrak{O}^c} \mathcal{M}^c.$$

Proof. Choose an $\mathfrak{O}$-basis $e_1, \ldots, e_d$ of $\mathcal{M}$, and let $A \in \mathrm{M}_d(\mathfrak{O})$ be the corresponding matrix of Frobenius, so $(\varphi(e_1), \ldots, \varphi(e_d)) = (e_1, \ldots, e_d)A$. We write $\mathfrak{e}_i := 1 \otimes e_i \in \varphi^*\mathcal{M}$ for the induced basis of $\varphi^*\mathcal{M}$. Using the definition of $\mathrm{Fil}^i\varphi^*\mathcal{M}$, it is not difficult to see that there is a matrix $B \in \mathrm{M}_d(\mathfrak{O})$ satisfying $AB = BA = E^r I_d$ and with the property that $\mathrm{Fil}^r\varphi^*\mathcal{M}$ is generated by $(\alpha_1, \ldots, \alpha_d) := (\mathfrak{e}_1, \ldots, \mathfrak{e}_d)B$. As promised, we denote by $\mathfrak{e}_i^c$, A^c, *etc.* the objects in the classical setting corresponding to a choice $e_1^c, \ldots, e_d^c$ of $\mathfrak{O}^c$-basis of $\mathcal{M}^c$. Let $X \in \mathrm{GL}_d(\widetilde{B}_\alpha)$ be the matrix determined by the requirement $\widetilde{\xi}(\mathfrak{e}_1, \ldots, \mathfrak{e}_d) = (\mathfrak{e}_1^c, \ldots, \mathfrak{e}_d^c)X$. As $\widetilde{\xi}$ is compatible with both Frobenius and filtrations, we find that

$$\widetilde{\xi}\circ\varphi(\mathfrak{e}_1, \ldots, \mathfrak{e}_d) = (\mathfrak{e}_1^c, \ldots, \mathfrak{e}_d^c)X\varphi(A) = \varphi\circ\widetilde{\xi}(\mathfrak{e}_1, \ldots, \mathfrak{e}_d) = (\mathfrak{e}_1^c \ldots, \mathfrak{e}_d^c)\varphi(A^c)\varphi(X)$$

and there exists a matrix $Y \in \mathrm{GL}_d(\widetilde{B}_\alpha)$ with

$$\widetilde{\xi}(\alpha_1, \ldots, \alpha_d) = (\alpha_1^c, \ldots, \alpha_d^c)Y.$$

We conclude that $X\varphi(A) = \varphi(A^c)\varphi(X)$ and $XB = B^cY$. Since $\widetilde{B}_\alpha$ is an integral domain, the facts that $B = E(u)^r A^{-1}$ and $B^c = E(u^c)^r (A^c)^{-1}$ imply that $A^c X E(u)^r = E(u^c)^r Y A$. Due to Lemma 2.1.3, both $E(u)$ and $E(u^c)$ are generators of $\mathrm{Fil}^1 W(R)_F$, so $\mu := E(u^c)/E(u)$ is a unit in $W(R)_F$. We therefore have the relation $A^c X = \mu^r Y A$. Combining this with the equality $X\varphi(A) = \varphi(A^c)\varphi(X)$ yields $X = \varphi(\mu^r Y)$, and we deduce $A^c\varphi(\mu^r Y) = \mu^r Y A$. Defining a $\widetilde{B}_\alpha$-linear map

$$\eta : \widetilde{B}_\alpha \otimes_{\mathfrak{O}} \mathcal{M} \longrightarrow \widetilde{B}_\alpha \otimes_{\mathfrak{O}^c} \mathcal{M}^c$$

by the requirement $(\eta(e_1), \ldots, \eta(e_d)) = (e_1^c, \ldots, e_d^c)\mu^r Y$, one then checks that η provides the claimed Frobenius-compatible isomorphism. □

Recall that Lemma 2.2.1 gives inclusions $\mathfrak{O} \subset \widetilde{B}^+_{\mathrm{rig},F} \subset \widetilde{B}_\alpha$.

Corollary 4.4.3. *The isomorphism η of Lemma* 4.4.2 *descends to a $\widetilde{B}^+_{\mathrm{rig},F}$-linear and Frobenius-compatible isomorphism*

$$\eta_{\mathrm{rig}} : \widetilde{B}^+_{\mathrm{rig},F} \otimes_{\mathfrak{O}} \mathcal{M} \simeq \widetilde{B}^+_{\mathrm{rig},F} \otimes_{\mathfrak{O}^c} \mathcal{M}^c.$$

Proof. We will use the notation of the proof of Lemma 4.4.2. Let us put $Z := \mu^r Y \in \mathrm{GL}_d(\widetilde{B}_\alpha)$, so that $\eta(e_1, \ldots, e_d) = (e_1, \ldots, e_d)Z$, and note that $A^c\varphi(Z) = ZA$ as η is compatible with Frobenius. To prove the corollary, it suffices to show that both Z and Z^{-1} have entries in $\widetilde{B}^+_{\mathrm{rig},F}$. We will show that

$Z \in \mathrm{M}_d(\widetilde{B}^+_{\mathrm{rig},F})$; the proof of the corresponding fact for Z^{-1} is similar and is left to the reader. By the definition of $\widetilde{B}^+_{\mathrm{rig},F}$, it suffices to show that for any m, there exists $Z_m \in \mathrm{M}_d(B^+_{\mathrm{max},F})$ with $\varphi^m(Z_m) = Z$, which we prove by induction on m. The base case $m = 0$ is obvious, as $\widetilde{B}_\alpha \subset B^+_{\mathrm{max},F}$. Now suppose that Z_m exists, and note that from the equality $A^c\varphi(Z) = ZA$ we obtain $E(u)^r Z = A^c\varphi(Z)B$. We may write $A^c = \varphi^{m+1}(A_{m+1})$ and $B = \varphi^{m+1}(B_{m+1})$ thanks to Lemma 2.2.1, and we then have $E(u)^r Z = \varphi^{m+1}(A_{m+1}Z_mB_{m+1})$. Finally, Lemma 2.1.5 implies the existence of Z_{m+1}. □

We can now prove Theorem 4.4.1:

Proof of Theorem 4.4.1. We first prove that $\mathcal{M}$ is étale, and to do so we will freely use the results and notation of [Ked04]. By the main theorem of [Ked04], $\mathcal{M}$ is étale if and only if $\mathcal{M}$ is *pure of slope 0*. Hence $\mathcal{M}^c$ is pure of slope 0 thanks Theorem 4.3.1 and our remarks immediately following Theorem 4.4.1. Since the slope filtration of $\mathcal{M}$ does not change after tensoring with the ring $\Gamma^{\mathrm{alg}}_{\mathrm{an,con}}$ constructed in [Ked04], it is enough to show that

$$\Gamma^{\mathrm{alg}}_{\mathrm{an,con}} \otimes_{\mathfrak{O}} \mathcal{M} \simeq \Gamma^{\mathrm{alg}}_{\mathrm{an,con}} \otimes_{\mathfrak{O}^c} \mathcal{M}^c.$$

as φ-modules over $\Gamma^{\mathrm{alg}}_{\mathrm{an,con}}$, and to do this it is enough thanks to Lemma 4.4.2 to prove that $\widetilde{B}^+_{\mathrm{rig},F} \subset \Gamma^{\mathrm{alg}}_{\mathrm{an,con}}$. But this follows from Berger's construction [Ber02, §2.3] (strictly speaking, [Ber02, §2.3] deals only with the case $F = \mathbb{Q}_p$, but see the last paragraph of [Ber14, §3] for the general case. We also warn the readers that Berger use $\widetilde{\mathbf{B}}^\dagger_{\mathrm{rig}}$ to denote $\widetilde{B}^\dagger_{\mathrm{rig},F}$ in this paper, while his $\widetilde{\mathbf{B}}^\dagger_{\mathrm{rig},F}$ means a different ring from ours), as he proves that $\widetilde{B}^+_{\mathrm{rig},F} \subset B^\dagger_{\mathrm{rig},F} = \widetilde{\mathbf{B}}^\dagger_{\mathrm{rig}} = \Gamma^{\mathrm{alg}}_{\mathrm{an,con}}$ (see the table over [Ber08, Prop. 1.1.12] for a helpful comparison of the various notations used by different authors). It follows that $\mathcal{M}$ is étale.

Now the proof that $\mathcal{M} := \mathcal{M}(D)$ admits a descent to a Kisin module $\mathfrak{M}$ is exactly the same as the proof of Lemma 1.3.13 in [Kis06], so we just sketch the highlights. As $\mathcal{M}$ is étale, there exists a finite free $\mathcal{O}_{\mathcal{R}^b}$-module $\mathcal{N}$ with Frobenius endomorphism $\varphi_{\mathcal{N}}$ satisfying

$$\mathcal{R} \otimes_{\mathcal{O}_{\mathcal{R}^b}} \mathcal{N} \simeq \mathcal{R} \otimes_{\mathfrak{O}} \mathcal{M} =: \mathcal{M}_{\mathcal{R}} \tag{4.4.2}$$

Proposition 6.5 in [Ked04] shows that it is possible to select an $\mathcal{R}$-basis of $\mathcal{M}_{\mathcal{R}}$ whose $\mathcal{R}^b$-span is exactly $\mathcal{N}[1/p]$ and whose $\mathfrak{O}$-span is $\mathcal{M}$ via the identifications (4.4.2). Define $\mathcal{M}^b \subseteq \mathcal{M}$ to be the $\mathfrak{S}_F[1/p]$-span of this basis. The equality $\mathfrak{S}_F[1/p] = \mathcal{R}^b \cap \mathfrak{O}$ provides the intrinsic description $\mathcal{M}^b = \mathcal{M} \cap \mathcal{N}[1/p]$; in particular, $\mathcal{M}^b$ is φ-stable and of E-height r. Let $\mathfrak{M}' := \mathcal{M}^b \cap \mathcal{N}$ and put $\mathfrak{M} := (\mathcal{O}_{\mathcal{R}^b} \otimes_{\mathfrak{S}_F} \mathfrak{M}') \cap \mathfrak{M}'[1/p] \subset \mathcal{N}[1/p]$. Then $\mathfrak{M}$ is a finite and φ-stable $\mathfrak{S}_F$-submodule of $\mathcal{N}[1/p]$. It follows from the structure theorem of finite $\mathfrak{S}_F$-modules [Fon90, Proposition 1.2.4] that $\mathfrak{M}$ is in fact finite and *free* over $\mathfrak{S}_F$. To see that $\mathfrak{M}$ has E-height r, it suffices to check that $\det(\varphi_{\mathfrak{M}}) = E^s w$ for some $w \in \mathfrak{S}_F^\times$. But $\mathfrak{M}[1/p] = \mathcal{M}^b$ and $\mathcal{M}^b$ is of finite E-height, so $\det(\varphi_{\mathfrak{M}}) = p^m E^s w$

for some $w \in \mathfrak{S}_F^\times$; as $\mathcal{M}$ is pure of slope 0 (equivalently, $\det(\varphi_{\mathfrak{M}}) \in \mathcal{O}_{\mathcal{R}^b}^\times$), we must in fact have $m = 0$. □

4.5. Full-faithfulness of restriction. Fix an object V of $\mathrm{Rep}_F^{F\text{-cris},r}(G)$, and let $\mathfrak{M}$ be a Kisin module associated to V via Theorem 4.4.1 (2).

Proposition 4.5.1. *There exists a natural $\mathcal{O}_F$-linear injection*

$$\iota : T_{\mathfrak{S}_F}(\mathfrak{M}) \hookrightarrow V \simeq V_{\mathrm{cris},F}(D)$$

that is moreover $G_{\underline{\pi}}$-equivariant. In particular, $V_{\mathfrak{S}}(\mathfrak{M}) \simeq V$ as $F[G_{\underline{\pi}}]$-modules.

Proof. Set $\mathcal{M} := \mathcal{M}(D)$. As $\mathcal{M} \simeq \mathfrak{O} \otimes_{\mathfrak{S}_F} \mathfrak{M}$, we have a natural injection

$$\iota' : T_{\mathfrak{S}_F}(\mathfrak{M}) = \mathrm{Hom}_{\mathfrak{S}_F,\varphi}(\mathfrak{M}, \mathfrak{S}_F^{\mathrm{ur}}) \hookrightarrow \mathrm{Hom}_{\mathfrak{O},\varphi,\mathrm{Fil}}(\varphi^*\mathcal{M}, B_{\mathrm{cris},F}^+),$$

uniquely determined by the requirement that for any $h : \mathfrak{M} \to \mathfrak{S}_F^{\mathrm{ur}}$, the value of $\iota'(h)$ on any simple tensor $s \otimes m \in \mathfrak{O} \otimes_{\varphi,\mathfrak{S}_F} \mathfrak{M} \simeq \varphi^*\mathcal{M}$ is given by

$$\iota'(h)(s \otimes m) = s\varphi(h(m)).$$

Using the fact that $E \in \mathrm{Fil}^1 W(R)_F$, one checks that this really does define a filtration-compatible $\mathfrak{O}$-linear homomorphism $\iota'(h) : \varphi^*\mathcal{M} \to B_{\mathrm{cris},F}^+$.

On the other hand, the isomorphism $\xi_\alpha : \mathfrak{O}_\alpha \otimes_{F_0} D \xrightarrow{\sim} \mathfrak{O}_\alpha \otimes_{\mathfrak{O}} \varphi^*\mathcal{M}$ of Lemma 4.2.2 induces, thanks to Lemma 4.2.3, a natural injection

$$\begin{tikzcd} \mathrm{Hom}_{\mathfrak{O},\varphi,\mathrm{Fil}}(\varphi^*\mathcal{M}, B_{\mathrm{cris},F}^+) \arrow[r, hook, "h\mapsto 1\otimes h"] & \mathrm{Hom}_{\mathfrak{O}_\alpha,\varphi,\mathrm{Fil}}(\mathfrak{O}_\alpha \otimes_{\mathfrak{O}} \varphi^*\mathcal{M}, B_{\mathrm{cris},F}^+) \arrow[d, "h\mapsto h\circ\xi_\alpha", "\simeq"'] \\ & \mathrm{Hom}_{\mathfrak{O}_\alpha,\varphi,\mathrm{Fil}}(\mathfrak{O}_\alpha \otimes_{F_0} D, B_{\mathrm{cris},F}^+) \end{tikzcd}$$

and we claim that $\mathrm{Hom}_{\mathfrak{O}_\alpha,\varphi,\mathrm{Fil}}(\mathfrak{O}_\alpha \otimes_{F_0} D, B_{\mathrm{cris},F}^+) = V_{\mathrm{cris},F}(D)$. By definition,

$$V_{\mathrm{cris},F}(D) = \mathrm{Hom}_{F_0,\varphi}(D, B_{\mathrm{cris},F}^+) \cap \mathrm{Hom}_{K,\mathrm{Fil}}(D_{F_0,K}, B_{\mathrm{dR}}^+),$$

and it is clear that $\mathrm{Hom}_{\mathfrak{O}_\alpha,\varphi}(\mathfrak{O}_\alpha \otimes_{F_0} D, B_{\mathrm{cris},F}^+) = \mathrm{Hom}_{F_0,\varphi}(D, B_{\mathrm{cris},F}^+)$. Since the injection $\mathfrak{O}_\alpha \otimes_{\mathfrak{O}} D \hookrightarrow \widehat{\mathfrak{S}}_0 \otimes D_{F_0,K}$ is compatible with filtrations by the very construction of the filtration on $\mathfrak{O}_\alpha \otimes_{F_0} D$, we conclude that

$$\begin{aligned} &\mathrm{Hom}_{F_0}(D, B_{\mathrm{cris},F}^+) \cap \mathrm{Hom}_{\mathfrak{O}_\alpha,\mathrm{Fil}}(\mathfrak{O}_\alpha \otimes_{F_0} D, B_{\mathrm{cris},F}^+) \\ &\quad \simeq \mathrm{Hom}_{F_0}(D, B_{\mathrm{cris},F}^+) \cap \mathrm{Hom}_{K,\mathrm{Fil}}(D_{F_0,K}, B_{\mathrm{dR}}^+), \end{aligned}$$

which gives our claim.

We thus obtain a natural injection $\iota : T_{\mathfrak{S}_F}(\mathfrak{M}) \hookrightarrow V_{\mathrm{cris},F}(D)$ which is visibly compatible with the given $G_{\underline{\pi}}$-actions. □

Combining Theorem 4.4.1 and Proposition 3.3.3, we deduce Theorem 1.0.1:

Corollary 4.5.2. *Let V be an object of $\mathrm{Rep}_F^{F\text{-cris},r}(G)$ and $T \subset V$ a $G_{\underline{\pi}}$-stable $\mathcal{O}_F$-lattice. Then there is a Kisin module $\mathfrak{M}$ of E-height r with $T_{\mathfrak{S}_F}(\mathfrak{M}) \simeq T$.*

Remark 4.5.3. It is an important point that in our general setup, the Kisin module $\mathfrak{M}$ may *not* be unique for a given T, contrary to the classical situation. Indeed, the "cyclotomic case" of Example 2.1.2 is a prototypical instance of such non-uniqueness: let T be the trivial character, $\mathfrak{M} = \mathfrak{S}_F$ the trivial rank-1 Kisin module, and $\mathfrak{M}' = u\mathfrak{S}_F \subset \mathfrak{M}$. Since $\varphi(u) = E(u)u$, one sees that $\mathfrak{M}'$ is also a Kisin module and $T_{\mathfrak{S}_F}(\mathfrak{M}) = T_{\mathfrak{S}_F}(\mathfrak{M}') = T$.

We now prove Theorem 1.0.2:

Theorem 4.5.4. *Assume that $\varphi^n(f(u)/u)$ is not a power of $E(u)$ for any $n \geq 0$ and that $v_F(a_1) > r$. Then the restriction functor $\mathrm{Rep}_F^{F\text{-cris},r}(G) \to \mathrm{Rep}_F(G_{\underline{\pi}})$ given by $V \rightsquigarrow V\big|_{G_{\underline{\pi}}}$ is fully faithful.*

Remark 4.5.5. We suspect that the theorem remains valid if we drop the assumption that $\varphi^n(f(u)/u)$ is not a power of E for any $n \geq 0$. When $r = 1$, we will show that this is indeed the case in the next section.

In order to prove Theorem 4.5.4, we prepare several preliminaries. In what follows, we keep our running notation with $f(u) = u^p + a_{p-1}u^{p-1} + \cdots + a_1 u$, and we assume throughout that $\varpi^{r+1} | a_1$ in $\mathcal{O}_F$.

Let $\mathfrak{M} \in \mathrm{Mod}_{\mathfrak{S}_F}^{\varphi,r}$ and set $M := \varphi^*\mathfrak{M}/u\varphi^*\mathfrak{M}$.

Lemma 4.5.6. *There exists a unique φ-equivariant isomorphism*

$$\xi_\alpha : \mathfrak{O}_\alpha \otimes_{\mathcal{O}_{F_0}} M \xrightarrow{\ \simeq\ } \mathfrak{O}_\alpha \otimes_{\mathfrak{S}_F} \varphi^*\mathfrak{M}$$

whose reduction modulo u is the identity on M.

Proof. The proof is similar to that of Proposition 2.4.1 in [Liu11], and is motivated by the proof of Lemma 1.2.6 in [Kis06]. Choose an $\mathfrak{S}_F$-basis $\mathfrak{e}_1, \ldots, \mathfrak{e}_d$ of $\mathfrak{M}$ and let $A \in \mathrm{M}_d(\mathfrak{S}_F)$ be the resulting matrix of φ; *i.e.* $(\varphi(\mathfrak{e}_1), \ldots, \varphi(\mathfrak{e}_d)) = (\mathfrak{e}_1, \ldots, \mathfrak{e}_d)A$. Then $e_i := 1 \otimes \mathfrak{e}_i$ forms a basis of $\varphi^*\mathfrak{M}$ and we have $(\varphi(e_1), \ldots, \varphi(e_d)) = (e_1, \ldots, e_d)\varphi(A)$. Put $A_0 := A \bmod u$ and $\bar{e}_i := e_i \bmod u$. Then we have $(\varphi(\bar{e}_1), \ldots, \varphi(\bar{e}_d)) = (\bar{e}_1, \ldots, \bar{e}_d)\varphi(A_0)$. If the map ξ_α of the Lemma exists, then writing $f_i := \xi_\alpha(\bar{e}_i) \in \mathfrak{O}_\alpha \otimes_{\mathfrak{S}_F} \varphi^*\mathfrak{M}$ and denoting by $Y \in \mathrm{GL}_d(\mathfrak{O}_\alpha)$ the matrix with $(f_1, \ldots, f_d) = (e_1, \ldots, e_d)Y$, we necessarily have $Y \equiv I_d \bmod u$ and

$$Y\varphi(A_0) = \varphi(A)\varphi(Y). \tag{4.5.1}$$

Conversely, if (4.5.1) has a solution $Y \in \mathrm{GL}_d(\mathfrak{O}_\alpha)$ satisfying $Y \equiv I_d \bmod u$, then we may define ξ_α by $\xi_\alpha(\bar{e}_1, \ldots, \bar{e}_d) = (e_1, \ldots, e_d)Y$. Thus, it remains to solve Equation (4.5.1). Put

$$Y_n := \varphi(A) \cdots \varphi^n(A)\varphi^n(A_0^{-1}) \cdots \varphi(A_0^{-1}). \tag{4.5.2}$$

We claim that the sequence $\{Y_n\}_n$ converges to a matrix $Y \in \mathrm{M}_d(\mathfrak{O}_\alpha)$. To see this, note that there exists $B_0 \in \mathrm{GL}_d(\mathfrak{O}_\alpha)$ with $A_0 B_0 = \varpi^r I_d$ since $\mathfrak{M}$ has height r. It follows that $AA_0^{-1} = I_d + \frac{u}{\varpi^r}Z$ for $Z \in \mathrm{M}_d(\mathfrak{S}_F)$. Thus,

$$Y_n = Y_{n-1} + \varphi(A) \cdots \varphi^{n-1}(A)\frac{\varphi^n(u)}{\varpi^{rn}}\varphi^n(Z)\varphi^{n-1}(B_0) \ldots \varphi(B_0),$$

so to prove our claim it suffices to show that $\varphi^n(u)/\varpi^{rn}$ converges to 0 in $\mathfrak{O}_\alpha$, which is the content of Lemma 2.2.2.

To prove that Y is invertible, we compute its determinant. Put $d := \operatorname{rank}_{\mathfrak{S}_F}\mathfrak{M}$, and observe that since $\wedge^d\mathfrak{M}$ has finite E-height, we have $\det(A) = \gamma E^m$ for some $\gamma \in \mathfrak{S}_F^\times$. It follows that $\det(\varphi(A)\varphi(A_0^{-1})) = \gamma'(\frac{\varphi(E)}{\varpi})^m$ for some $\gamma' \in \mathfrak{S}_F^\times$. One then checks that $\varphi(E)/\varpi$ is a unit in $\mathfrak{O}_\alpha$, and hence that $\det(Y)$ is a unit in $\mathfrak{O}_\alpha$ so Y is invertible as desired.

Finally, we prove that the solution Y to (4.5.1) that we have constructed is unique. Suppose that equation (4.5.1) admits two solutions Y, Y' satisfying $Y,\ Y' \equiv I_d \bmod u$. Then their difference is also a solution $Y - Y' = uZ$ for $Z \in \mathrm{M}_d(\mathfrak{O}_\alpha)$. Equation (4.5.1) then implies that for all n we have

$$\begin{aligned} Y - Y' = \varphi(A)\varphi(Y)\varphi(A_0^{-1}) &= \varphi(A)\cdots\varphi^n(A)\varphi^n(Y)\varphi^n(A_0^{-1})\cdots\varphi(A_0^{-1}) \\ &= \varphi(A)\cdots\varphi^n(A)\frac{\varphi^n(u)}{\varpi^{rn}}\varphi^n(Z)\varphi^n(B_0)\cdots\varphi(B_0) \end{aligned}$$

As $\varphi^n(u)/\varpi^{rn}$ converges to 0 in $\mathfrak{O}_\alpha$, we conclude that $Y = Y'$ as desired. □

For $\mathfrak{M} \in \mathrm{Mod}_{\mathfrak{S}_F}^{\varphi,r}$, let us write $D(\mathfrak{M}) = \xi_\alpha(M[1/p]) \subset \mathfrak{O}_\alpha \otimes_{\mathfrak{S}_F} \varphi^*\mathfrak{M}$ for the image of $M[1/p]$ under the map of Lemma 4.5.6. If $\mathfrak{M}$ is a Kisin module associated to some F-crystalline G-representation V with $D := D_{\mathrm{cris},F}(V)$ (*i.e.* $\mathfrak{O} \otimes_{\mathfrak{S}_F} \mathfrak{M} \simeq \mathcal{M}(D)$), then by the very construction of $\mathcal{M}(D)$ there is a natural φ-compatible inclusion $D \simeq \varphi^*D \hookrightarrow \varphi^*\mathcal{M}(D)$ which, thanks to Lemma 4.2.2, becomes an isomorphism after tensoring over $\mathfrak{O}$ with $\mathfrak{O}_\alpha$. Recalling that $\mathfrak{O}_\alpha \otimes_{\mathfrak{S}_F} \mathfrak{M} \simeq \mathcal{M}(D)$, we therefore have a φ-equivariant inclusion

$$D \xhookrightarrow{d\mapsto 1\otimes d} \mathfrak{O}_\alpha \otimes_{F_0} D \xrightarrow[\simeq]{4.2.2} \mathfrak{O}_\alpha \otimes_{\mathfrak{O}} \varphi^*\mathcal{M}(D) \simeq \mathfrak{O}_\alpha \otimes_{\mathfrak{S}_F} \varphi^*\mathfrak{M} \tag{4.5.3}$$

via which we view D as a φ-stable F_0-subspace of $\mathfrak{O}_\alpha \otimes_{\mathfrak{S}_F} \varphi^*\mathfrak{M}$.

Corollary 4.5.7. *Let $V \in \mathrm{Rep}_F^{F\text{-cris},r}(G)$. If $\mathfrak{M} \in \mathrm{Mod}_{\mathfrak{S}_F}^{\varphi,r}$ is a Kisin module attached to $D := D_{\mathrm{cris},F}(V)$, then $D(\mathfrak{M}) = D$ inside $\mathfrak{O}_\alpha \otimes_{\mathfrak{S}} \varphi^*\mathfrak{M}$.*

Proof. The reduction of (4.5.3) modulo u is the φ-compatible isomorphism

$$D \simeq (\mathfrak{O}_\alpha \otimes_{\mathfrak{S}_F} \varphi^*\mathfrak{M}) \bmod u \simeq M[1/p].$$

Since the map ξ_α of Lemma 4.5.6 reduces to the identity modulo u, we conclude that *both* D and $D(\mathfrak{M})$ inside $\mathfrak{O}_\alpha \otimes_{\mathfrak{S}_F} \varphi^*\mathfrak{M}$ are φ-equivariant liftings of $M[1/p]$; the uniqueness aspect of Lemma 4.5.6 then forces $D = D(\mathfrak{M})$ as claimed. □

It follows from Corollary 4.5.7 that the map ξ_α of Lemma 4.5.6 coincides with that of Lemma 4.2.2, which justifies our notation.

Recall that $V_{\mathfrak{S}}(\mathfrak{M}) = T_{\mathfrak{S}}(\mathfrak{M})[1/p]$ for $\mathfrak{M} \in \mathrm{Mod}_{\mathfrak{S}_F}^{\varphi,r}$.

Lemma 4.5.8. *Let $\mathfrak{f} : \mathfrak{M} \to \mathfrak{M}'$ be any morphism of height-r Kisin modules, and let $\mathfrak{f}_\alpha$ be the scalar extension $\mathfrak{f}_\alpha : \mathfrak{O}_\alpha \otimes_{\mathfrak{S}_F} \varphi^*\mathfrak{M} \to \mathfrak{O}_\alpha \otimes_{\mathfrak{S}_F} \varphi^*\mathfrak{M}'$ of $\varphi^*\mathfrak{f}$ along $\mathfrak{S}_F \to \mathfrak{O}_\alpha$. Then $\mathfrak{f}_\alpha(D(\mathfrak{M})) \subset D(\mathfrak{M}')$.*

Proof. Put $V = V_{\mathfrak{S}}(\mathfrak{M})$ and $V' = V_{\mathfrak{S}}(\mathfrak{M}')$ and denote by $\psi = V_{\mathfrak{S}}(\mathfrak{f}) : V' \to V$ the induced map. By Proposition 3.3.3, we can modify $\mathfrak{M}$ and $\mathfrak{M}'$ (inside $\mathfrak{M}[1/p]$ and $\mathfrak{M}'[1/p]$, respectively) so that $\mathfrak{f}$ may be decomposed by two exact sequences inside $\mathrm{Mod}_{\mathfrak{S}_F}^{\varphi,r}$:

$$0 \longrightarrow \mathfrak{L} \longrightarrow \mathfrak{M} \xrightarrow{\mathfrak{f}} \mathfrak{N} \longrightarrow 0 \quad \text{and} \quad 0 \longrightarrow \mathfrak{N} \longrightarrow \mathfrak{M}' \longrightarrow \mathfrak{N}' \longrightarrow 0,$$

where $\mathfrak{N} = \mathfrak{f}(\mathfrak{M})$. From the construction of ξ_α in Lemma 4.5.6 (in particular, from the explicit construction of Y in (4.5.2)), we obtain the exact sequences

$$0 \longrightarrow D(\mathfrak{L}) \longrightarrow D(\mathfrak{M}) \xrightarrow{\mathfrak{f}_\alpha} D(\mathfrak{N}) \longrightarrow 0$$

and

$$0 \longrightarrow D(\mathfrak{N}) \longrightarrow D(\mathfrak{M}') \longrightarrow D(\mathfrak{N}') \longrightarrow 0 ,$$

which shows that $\mathfrak{f}_\alpha(D(\mathfrak{M})) \subset D(\mathfrak{M}')$ as claimed. $\square$

Proof of Theorem 4.5.4. Let V', V be two objects of $\mathrm{Rep}_F^{F\text{-cris},r}(G)$, and set $D = D_{\mathrm{cris},F}(V)$ and $D' = D_{\mathrm{cris},F}(V')$. Suppose that there exists an F-linear map $h : V_{\mathrm{cris},F}(D') \to V_{\mathrm{cris},F}(D)$ that is $G_{\underline{\pi}}$-equivariant. By Corollary 4.5.2, there exist $G_{\underline{\pi}}$-stable $\mathcal{O}_F$-lattices T and T' inside $V_{\mathrm{cris},F}(D)$ and $V_{\mathrm{cris},F}(D')$, respectively, with $h(T') \subseteq T$, and objects $\mathfrak{M}$ and $\mathfrak{M}'$ of $\mathrm{Mod}_{\mathfrak{S}_F}^{\varphi,r}$ such that $T_{\mathfrak{S}}(\mathfrak{M}) \simeq T$ and $T_{\mathfrak{S}}(\mathfrak{M}') \simeq T'$ via the map ι of Proposition 4.5.1. By Proposition 3.3.5, there exists a map $\mathfrak{f} : \mathfrak{M} \to \mathfrak{M}'$ in $\mathrm{Mod}_{\mathfrak{S}_F}^{\varphi,r}$ with $V_{\mathfrak{S}_F}(\mathfrak{f}) \simeq h$. We may therefore realize h as the composite

$$V_{\mathrm{cris},F}(D') \xrightarrow[\sim]{\iota^{-1}} V_{\mathfrak{S}}(\mathfrak{M}')[1/p] \xrightarrow{V_{\mathfrak{S}}(\mathfrak{f})} V_{\mathfrak{S}}(\mathfrak{M})[\tfrac{1}{p}] \xrightarrow[\sim]{\iota} V_{\mathrm{cris},F}(D) ,$$

where ι is constructed using the isomorphism $\xi'_\alpha : \mathfrak{O}_\alpha \otimes_{\mathfrak{S}_F} \varphi^*\mathfrak{M} \simeq \mathfrak{O}_\alpha \otimes D$ of Lemma 4.2.3. Due to Lemma 4.5.8 and Corollary 4.5.7, we know that $\mathfrak{f}$ induces a map $\mathfrak{f}_\alpha : \mathfrak{O}_\alpha \otimes_{\mathfrak{S}_F} \varphi^*\mathfrak{M} \to \mathfrak{O}_\alpha \otimes_{\mathfrak{S}_F} \varphi^*\mathfrak{M}'$ carrying D to D', so for $a \in V_{\mathrm{cris},F}(D') \subset \mathrm{Hom}_{F_0,\varphi}(D, B^+_{\mathrm{cris},F})$ we have $h(a) = a \circ \bar{\mathfrak{f}} \in V_{\mathrm{cris},F}(D') \subset \mathrm{Hom}_{F_0,\varphi}(D', B^+_{\mathrm{cris},F})$ where we write $\bar{\mathfrak{f}} : D \to D'$ for the map $\mathfrak{f}_\alpha|_D$. It follows at once that h is compatible with the action of $G = G_K$, as desired. $\square$

We note that Theorem 4.5.4 is false if we replace "$\mathrm{Rep}_F^{F\text{-cris}}(G)$" with "$\mathrm{Rep}_F^{\mathbb{Q}_p\text{-cris}}(G)$," as the following example shows:

Example 4.5.9. Consider again the setting of Example 2.1.2, with $K = \mathbb{Q}_p(\zeta_p)$, $\pi = \zeta_p - 1$ and $\varphi(u) = f(u) = (1+u)^p - 1$, and $K_{\underline{\pi}} = \bigcup_{n\geq 1} \mathbb{Q}_p(\zeta_{p^n})$. Let $F = \mathbb{Q}_p$. Then the assumption of Theorem 4.5.4 is not satisfied as $a_1 = p$, and the restriction functor $\mathrm{Rep}_F^{\mathbb{Q}_p\text{-cris}}(G) \to \mathrm{Rep}_F(G_\infty)$ induced by $V \rightsquigarrow V\big|_{G_{\underline{\pi}}}$ is visibly *not* fully faithful: letting χ denote the p-adic cyclotomic character, we have $\chi\big|_{G_{\underline{\pi}}} = 1\big|_{G_{\underline{\pi}}}$, but $\chi \not\simeq 1$ as G-representations. On the other hand, if $F = K$ then we easily check that the assumptions of Theorem 4.5.4 are satisfied. Of course, there is no contradiction here as χ is *not* an F-crystalline representation because $\mathrm{HT}_\tau(\chi) = 1$ for *all* τ.

5. F-Barsotti-Tate groups

Recall that by an *F-Barsotti–Tate group* over $\mathcal{O}_K$, we mean a Barsotti–Tate group over $\mathcal{O}_K$ whose p-adic Tate module is an F-crystalline representation of $G := G_K$. In this section, we prove that the category of F-Barsotti–Tate groups over $\mathcal{O}_K$ is (anti)equivalent to the category of height-1 Kisin modules:

Theorem 5.0.10. *Assume $v_F(a_1) > 1$. Then there is an (anti)equivalence of categories between the category of Kisin modules of height 1 and the category of F-Barsotti-Tate groups.*

Using well-known results of Breuil, Kisin, Raynaud, and Tate, one shows as in [Liu13b, Theorem 2.2.1] that the p-adic Tate module gives an equivalence between thet category of F-Barsotti–Tate groups over $\mathcal{O}_K$ and the category $\mathrm{Rep}_{\mathcal{O}_F}^{F\text{-cris},1}(G)$ of G-stable $\mathcal{O}_F$-lattices inside F-crystalline representations with Hodge-Tate weights in $\{0,1\}$. Thus, to prove Theorem 5.0.10 we must construct an (anti)equivalence between $\mathrm{Mod}_{\mathfrak{S}_F}^{\varphi,1}$ and $\mathrm{Rep}_{\mathcal{O}_F}^{F\text{-cris},1}(G)$. In what follows, we show that for each $\mathfrak{M} \in \mathrm{Mod}_{\mathfrak{S}_F}^{\varphi,1}$ the natural $G_{\underline{\pi}}$-action on $T_{\mathfrak{S}}(\mathfrak{M})$ can be functorially extended to to a G-action such that $T_{\mathfrak{S}}(\mathfrak{M}) \in \mathrm{Rep}_{\mathcal{O}_F}^{F\text{-cris},1}(G)$. This construction will provide a contravariant functor $T_{\mathfrak{S}} : \mathrm{Mod}_{\mathfrak{S}_F}^{\varphi,1} \to \mathrm{Rep}_{\mathcal{O}_F}^{F\text{-cris},1}(G)$ that we will then prove is an (anti)equivalence.

5.1. A natural G-action on $T_{\mathfrak{S}}(\mathfrak{M})$.

Fix a Kisin module $\mathfrak{M}$ of height 1. In this subsection, we will construct a natural G-action on $T_{\mathfrak{S}}(\mathfrak{M})$ which extends the given action of $G_{\underline{\pi}}$. The key input to this construction is:

Lemma 5.1.1. *There exists a unique $W(R)_F$-semilinear G-action on $W(R)_F \otimes_{\mathfrak{S}_F} \mathfrak{M}$ that commutes with φ and satisfies*

1. *If $g \in G_{\underline{\pi}}$ and $m \in \mathfrak{M}$ then $g(1 \otimes m) = 1 \otimes m$;*
2. *If $m \in \mathfrak{M}$ then $1 \otimes (g(m) - m) \in \mathfrak{t}_F I^+(W(R)_F \otimes_{\mathfrak{S}_F} \mathfrak{M})$.*

Here, we remind the reader that $\mathfrak{t}_F \in W(R)_F$, constructed in Example 3.3.2, satisfies $\varphi(\mathfrak{t}_F) = E\mathfrak{t}_F$ and $\mathfrak{t}_F \not\equiv 0 \bmod \varpi$.

Proof. Fix an $\mathfrak{S}_F$-basis $\mathfrak{e}_1, \ldots, \mathfrak{e}_d$ of $\mathfrak{M}$ and let A be the resulting matrix of Frobenius, so $(\varphi(\mathfrak{e}_1), \ldots, \varphi(\mathfrak{e}_d)) = (\mathfrak{e}_1, \ldots, \mathfrak{e}_d)A$. Supposing that the required G-action exists, for any $g \in G$ we have a matrix $X_g \in \mathrm{M}_d(W(R)_F)$ with $(g\mathfrak{e}_1, \ldots, g\mathfrak{e}_d) = (\mathfrak{e}_1, \ldots, \mathfrak{e}_d)X_g$, and the requirement that g and φ commute is equivalent the matrix equation

$$X_g g(A) = A\varphi(X_g). \tag{5.1.1}$$

We claim that for each $g \in G$, equation (5.1.1) has a *unique* solution X_g satisfying the condition $X_g - I_d \in \mathrm{M}_d(\mathfrak{t}_F I^+)$. Granting this for a moment, it is easy to see that the Lemma follows once we check that $g \mapsto X_g$ really defines an action of G, which is equivalent to the cocycle condition $X_\sigma \sigma(X_\tau) = X_{\sigma\tau}$ for all $\sigma, \tau \in G$. But it is clear that $X_\sigma\sigma(X_\tau)$ and $X_{\sigma\tau}$ are both solutions to $X\sigma\tau(A) = A\varphi(X)$, and the condition $X - I_d \in \mathrm{M}_d(\mathfrak{t}_F I^+)$ holds for $X = X_{\sigma\tau}$ by our claim and for $X = X_\sigma \sigma X_\tau$ thanks to Lemma 2.3.4. Thus, the uniqueness aspect of our claim gives $X_\sigma \sigma(X_\tau) = X_{\sigma\tau}$, as desired.

It remains to prove our claim. Let us first dispense with the uniqueness aspect. Suppose that for some $g \in G$, equation (5.1.1) has two solutions X_1, X_2 satisfying $X_i - I_d \in \mathrm{M}_d(\mathfrak{t}_F I^+)$ for $i = 1, 2$. Then their difference is a solution as well, and has the form $X_1 - X_2 = \mathfrak{t}_F Z$ for some $Z \in \mathrm{M}_d(I^+)$. Equation (5.1.1) then takes the shape

$$\mathfrak{t}_F Z g(A) = A\varphi(\mathfrak{t}_F Z), \tag{5.1.2}$$

and we will show that this forces $Z = 0$, giving uniqueness. First, writing $\bar{Z} := Z \bmod \varpi \in \mathrm{M}_d(R)$, we note that it suffices to prove that $\bar{Z} = 0$: indeed, if $Z = \varpi Z_1$ for some $Z_1 \in \mathrm{M}_d(W(R)_F)$, then $Z_1 \in \mathrm{M}_d(I^+)$ is another solution to (5.1.2), so boot-strapping the argument gives $Z \in \cap_{n \geq 1} \varpi^n W(R)_F = \{0\}$. Now since $\mathfrak{M}$ has height 1, there exists a matrix $B \in \mathrm{M}_d(\mathfrak{S}_F)$ with $AB = EI_d$. On the other hand, we have $\varphi(\mathfrak{t}_F) = E\mathfrak{t}_F$ as noted above, so it follows from (5.1.2) that there exists a matrix $C \in \mathrm{M}_d(W(R)_F)$ with $Z = A\varphi(Z)C$. Reducing modulo ϖ gives a matrix equation $\bar{Z} = \bar{A}\varphi(\bar{Z})\bar{C}$ in $\mathrm{M}_d(R)$. If $\bar{Z} \neq 0$, then there exists an entry z, say, of $\bar{Z}$ which has minimal valuation. On the other hand, as $Z \in \mathrm{M}_d(I^+)$, we must have $v_R(z) > 0$. But the minimal possible valuation of entries in $\bar{A}\varphi(\bar{Z})\bar{C}$ is $pv_R(z) > v_R(z)$, which is a contradiction. Thus $\bar{Z} = 0$, settling uniqueness.

Finally, let us prove the existence of X_g solving (5.1.1) for each $g \in G$. For ease of notation, put

$$P_n := A\varphi(A)\cdots\varphi^n(A) \text{ and } Q_n := \varphi^n(g(A^{-1}))\cdots\varphi(g(A^{-1}))g(A^{-1}) \tag{5.1.3}$$

and define $X_n := P_n Q_n$. It suffices to prove the following:

(1) $X_n \in \mathrm{M}_d(W(R)_F)$ for all n;
(2) $X_n - I_d \in \mathrm{M}_d(\mathfrak{t}_F I^+)$ for all n;
(3) X_n converges as $n \to \infty$.

For (1) and (2), we argue by induction on n. When $n = 0$, by definition we have $X_0 = P_0 Q_0 = Ag(A^{-1}) = g(g^{-1}(A)A^{-1})$. On the other hand, by Lemma 2.3.2, we may write $g^{-1}A = A + \varphi(\mathfrak{t}_F)C$ for some $C \in \mathrm{M}_d(I^+)$, which gives

$$g^{-1}(A)A^{-1} = I_d + \varphi(\mathfrak{t}_F)CA^{-1} = I_d + \mathfrak{t}_F CEA^{-1} = I_d + \mathfrak{t}_F CB$$

thus proving (1) and (2) in the base case $n = 0$.

Now suppose we have proved $X_n = I_d + \mathfrak{t}_F C_n$ with $C_n \in \mathrm{M}_d(I^+)$, and let us show that X_{n+1} satisfies the same equation for some $C_{n+1} \in \mathrm{M}_d(I^+)$. Writing $Ag(A^{-1}) = I_d + \mathfrak{t}_F C_0$ with $C_0 \in \mathrm{M}_d(I^+)$, we have

$$X_{n+1} = X_n + P_n \varphi^{n+1}(\mathfrak{t}_F)\varphi^{n+1}(C_0)Q_n$$

Now $Eg(A^{-1}) \in \mathrm{M}_d(W(R)_F)$ as $g(E) = \mu_g E$ for some unit $\mu_g \in W(R)_F$, and we have $\varphi^{n+1}(\mathfrak{t}_F) = \varphi^n(E)\cdots\varphi(E)E\mathfrak{t}_F$. We conclude that the matrix $\widetilde{Q}_n := \varphi^{n+1}(\mathfrak{t}_F)Q_n$ lies in $\mathrm{M}_d(\mathfrak{t}_F W(R)_F)$, which gives $X_{n+1} \in \mathrm{M}_d(W(R)_F)$

and $X_{n+1} - I_d \in \mathrm{M}_d(\mathfrak{t}_F I^+)$ as desired. By construction, we then have

$$X_{n+1} - X_n = P_n \cdot \varphi^{n+1}(C_0) \cdot \widetilde{Q}_n,$$

with $P_n, \widetilde{Q}_n \in \mathrm{M}_d(W(R)_F)$ and since $\varphi^{n+1}(C_0)$ converges to 0 in $W(R)_F$, we conclude that X_n converges, which gives (3) and completes the proof. $\square$

Corollary 5.1.2. *The natural $G_{\underline{\pi}}$-action on $T_{\mathfrak{S}}(\mathfrak{M})$ can be functorially extended to an action of G. In particular, $T_{\mathfrak{S}}$ extends to a contravariant functor from $\mathrm{Mod}_{\mathfrak{S}_F}^{\varphi,1}$ to $\mathrm{Rep}_{\mathcal{O}_F}(G)$.*

Proof. By Lemma 3.3.1 (2), we have isomorphisms of $\mathcal{O}_F[G_{\underline{\pi}}]$-modules

$$\begin{aligned} T_{\mathfrak{S}}(\mathfrak{M}) &\simeq \mathrm{Hom}_{\mathfrak{S}_F,\varphi}(\mathfrak{M}, W(R)_F) \\ &\simeq \mathrm{Hom}_{W(R)_F,\varphi}(W(R)_F \otimes_{\mathfrak{S}_F} \mathfrak{M}, W(R)_F). \end{aligned} \tag{5.1.4}$$

Thanks to Lemma 5.1.1, we have an action of G on $W(R)_F \otimes_{\mathfrak{S}_F} \mathfrak{M}$ that extends the given action of $G_{\underline{\pi}}$, so the final term in (5.1.4) has an action of G given by

$$(g \circ h)(x) = g(h(g^{-1}(x))),\ \forall g \in G,\ \forall h \in \mathrm{Hom}_{W(R)_F,\varphi}(W(R)_F \otimes_{\mathfrak{S}_F} \mathfrak{M}, W(R)_F)$$

and one checks easily that this action extends the given action of $G_{\underline{\pi}}$ on $T_{\mathfrak{S}}(\mathfrak{M})$. It remains to prove that $T_{\mathfrak{S}}$ is a functor. So suppose that $h : \mathfrak{M} \to \mathfrak{M}'$ is a map in $\mathrm{Mod}_{\mathfrak{S}_F}^{\varphi,1}$ and let us check that the induced map $T_{\mathfrak{S}}(h) : T_{\mathfrak{S}}(\mathfrak{M}') \to T_{\mathfrak{S}}(\mathfrak{M})$ is indeed a map of $\mathcal{O}_F[G]$-modules. To do this, using (5.1.4), it suffices to show that the map

$$1 \otimes h : W(R)_F \otimes_{\mathfrak{S}_F} \mathfrak{M} \to W(R)_F \otimes_{\mathfrak{S}_F} \mathfrak{M}'$$

is G-equivariant, *i.e.* that $(1 \otimes h) \circ g = g \circ (1 \otimes h)$ for all $g \in G$. Choose $\mathfrak{S}_F$-bases $\mathfrak{e}_1, \ldots, \mathfrak{e}_d$ and $\mathfrak{e}'_1, \ldots, \mathfrak{e}'_{d'}$ of $\mathfrak{M}$ and $\mathfrak{M}'$, respectively, and let A and A' be the corresponding matrices of Frobenius, so $(\varphi(\mathfrak{e}_1), \ldots, \varphi(\mathfrak{e}_d)) = (\mathfrak{e}_1, \ldots, \mathfrak{e}_d)A$ and $(\varphi(\mathfrak{e}'_1), \ldots, \varphi(\mathfrak{e}'_{d'})) = (\mathfrak{e}'_1, \ldots, \mathfrak{e}'_{d'})A'$. Letting Z be the $d' \times d$-matrix with entries in $\mathfrak{S}_F$ determined by the relation $h(\mathfrak{e}_1, \ldots, \mathfrak{e}_d) = (\mathfrak{e}'_1, \ldots, \mathfrak{e}'_{d'})Z$, we seek to prove that $g \circ (1 \otimes h)(\mathfrak{e}_1, \ldots, \mathfrak{e}_d) = (1 \otimes h) \circ g(\mathfrak{e}_1, \ldots, \mathfrak{e}_d)$, which is equivalent to the matrix equation

$$X'_g g(Z) = Z X_g,$$

where X_g (resp. X'_g) is the matrix constructed in the proof of Lemma 5.1.1 giving the action of g on $\mathfrak{M}$ (resp. $\mathfrak{M}'$). By construction, $X_g = \lim_{n\to\infty} X_n$, and similarly for X'_g, so it suffices to check that $X'_n g(Z) = Z X_n$ for all n. From the very definition of $X_n = P_n Q_n$ and $X'_n = P'_n Q'_n$ via (5.1.3), this amounts to

$$P'_n Q'_n g(Z) = Z A P_n Q_n. \tag{5.1.5}$$

But as $\varphi \circ h = h \circ \varphi$, we have $A'\varphi(Z) = ZA$, or equivalently, $A'^{-1} Z = \varphi(Z) A^{-1}$, and the truth of equation (5.1.5) follows easily from the definition (5.1.3). $\square$

5.2. An equivalence of categories. In this subsection, we prove Theorem 5.0.10. Let us first recall the setup and some notation. For $\mathfrak{M} \in \mathrm{Mod}^{\varphi,1}_{\mathfrak{S}_F}$, put $\mathcal{M} := \mathfrak{O} \otimes_{\mathfrak{S}_F} \mathfrak{M}$ and define a decreasing filtration on $\varphi^*\mathcal{M}$ as in (4.2.2). Since $\mathfrak{M}$ has height 1, we have $\mathrm{Fil}^i\varphi^*\mathcal{M} = E^{i-1}\mathrm{Fil}^1\varphi^*\mathcal{M}$ for $i \geq 2$. Recall that we set $M := \varphi^*\mathfrak{M}/u\varphi^*\mathfrak{M}$ and let us put $D := D(\mathfrak{M}) := \xi_\alpha(M[1/p])$, which is naturally a φ-stable F_0-subspace of $\mathfrak{O}_\alpha \otimes_{\mathfrak{S}_F} \varphi^*\mathfrak{M}$ via the unique isomorphism ξ_α constructed in Lemma 4.5.6. Since $\xi_\alpha : \mathfrak{O}_\alpha \otimes_{F_0} D \to \mathfrak{O}_\alpha \otimes \varphi^*\mathcal{M}$ is an isomorphism, we may identify $\varphi^*\mathcal{M}/E\varphi^*\mathcal{M}$ with $D_{F_0,K} = K \otimes_{F_0} D$, and we write $\psi_\pi : \varphi^*\mathcal{M} \to D_{F_0,K}$ for the natural projection. We then define $\mathrm{Fil}^i D_{F_0,K} := \psi_\pi(\mathrm{Fil}^i\varphi^*\mathcal{M})$, and note that since $\mathrm{Fil}^2\varphi^*\mathcal{M} \subset E\varphi^*\mathcal{M}$, we have $\mathrm{Fil}^2 D_{F_0,K} = 0$. In this way we obtain from $\mathfrak{M}$ an object $D = D(\mathfrak{M})$ of $\mathrm{MF}^{\varphi}_{F_0,K}$. Suppose that $\mathfrak{M}' = \mathfrak{S}_F \cdot \mathfrak{e}$ is a rank-1 Kisin module with $\mathfrak{S}_F$-basis $\mathfrak{e}$. Then we have $\varphi(\mathfrak{e}) = \gamma E^m \mathfrak{e}$ with $\gamma \in \mathfrak{S}_F^\times$ a unit thanks to Example 3.3.2, and we call m the *minimal height* of $\mathfrak{M}'$.

Lemma 5.2.1. *With notation as above,*

(1) *The natural injection*

$$\mathfrak{O}_\alpha \otimes_{\mathfrak{O}} \varphi^*\mathcal{M} \xrightarrow{\xi'_\alpha} \mathfrak{O}_\alpha \otimes_{F_0} D \hookrightarrow \widehat{\mathfrak{S}}_0 \otimes D_{F_0,K}$$

is compatible with filtrations, where $\xi'_\alpha = (\xi_\alpha)^{-1}$.

(2) *Suppose $\mathfrak{M}$ has rank d. Then the minimal height of $\wedge^d\mathfrak{M}$ is $\dim_{K_0} \mathrm{Fil}^1 D_{F_0,K}$.*

Proof. Since $\mathrm{Fil}^i\varphi^*\mathcal{M} = E^{i-1}\mathrm{Fil}^1\varphi^*\mathcal{M}$ for $i \geq 2$, to prove (1) it suffices to check the given injection is compatible with Fil^1. As E is a generator of $\mathrm{Fil}^1\widehat{\mathfrak{S}}_0$, such compatibility is equivalent to the condition that $x \in \mathrm{Fil}^1\varphi^*\mathcal{M}$ if and only if $\psi_\pi(x) \in \mathrm{Fil}^1 D_{F_0,K}$. But this is clear as $E\varphi^*\mathcal{M} \subset \mathrm{Fil}^1\varphi^*\mathcal{M}$.

We now prove (2). Fix an $\mathfrak{S}_F$-basis $\mathfrak{e}_1, \ldots, \mathfrak{e}_d$ of $\mathfrak{M}$ and let $A \in \mathrm{M}_d(\mathfrak{S}_F)$ be the corresponding matrix of Frobenius. Since $\mathfrak{M}$ has height 1, there exists a matrix $B \in \mathrm{M}_d(\mathfrak{S}_F)$ with $AB = EI_d$. Defining $e_i = 1 \otimes \mathfrak{e}_i \in \varphi^*\mathcal{M}$, we easily check that $\{e_i\}$ is an $\mathfrak{O}$-basis of $\varphi^*\mathcal{M}$ with $(\alpha_1, \ldots, \alpha_d) := (e_1, \ldots, e_d)B$ an $\mathfrak{O}$-basis of $\mathrm{Fil}^1\varphi^*\mathcal{M}$.

Now the inclusion $\varphi^*\mathcal{M}/\mathrm{Fil}^1\varphi^*\mathcal{M} \subset \varphi^*\mathcal{M}/E\varphi^*\mathcal{M} = D_{F_0,K}$ realizes $\varphi^*\mathcal{M}/\mathrm{Fil}^1\varphi^*\mathcal{M}$ as a K-subspace of $D_{F_0,K}$, so there exists a basis $f_1, \ldots, f_d$ of $\varphi^*\mathcal{M}$ with the property that $f_1, \ldots f_s, Ef_{s+1}, \ldots, Ef_d$ generates $\mathrm{Fil}^1\varphi^*\mathcal{M}$. Since $\mathrm{Fil}^1 D_{F_0,K} = \psi_\pi(\mathrm{Fil}^1\varphi^*\mathcal{M})$ we have $\dim_K \mathrm{Fil}^1 D_{F_0,K} = s$. On the other hand, since $\alpha_1, \ldots, \alpha_d$ also generates $\mathrm{Fil}^1\varphi^*\mathcal{M}$, there exist invertible matrices $X, Y \in \mathrm{GL}_d(\mathfrak{O})$ with

$$B = X\Lambda Y \quad \text{for} \quad \Lambda = \mathrm{diag}(1, \ldots, 1, E, \ldots, E)$$

the diagonal matrix with s many 1's and $d - s$ many E's along the diagonal. Thus, $\det B = E^{d-s}\gamma$ for $\gamma \in \mathfrak{O}^\times$ a unit and since $AB = EI_d$ we then have $\det(A) = E^s\gamma^{-1}$. It follows that the minimal height of $\wedge^d\mathfrak{M}$ is $s = \dim_K \mathrm{Fil}^1 D_{F_0,K}$, as desired. □

Recall that we have defined $V_{\mathfrak{S}}(\mathfrak{M}) := F \otimes_{\mathcal{O}_F} T_{\mathfrak{S}}(\mathfrak{M})$.

Proposition 5.2.2. *With notation as above, we have $V_{\mathfrak{S}}(\mathfrak{M}) \simeq V_{\mathrm{cris},F}(D(\mathfrak{M}))$ as $F[G]$-modules. In particular, $V_{\mathfrak{S}}(\mathfrak{M})$ is crystalline with Hodge-Tate weights in $\{0,1\}$.*

Proof. The proof of Proposition 4.5.1 carries over *mutatis mutandis* to show that there exists a natural injection of $\mathcal{O}_F[G_{\underline{\pi}}]$-modules

$$\iota : T_{\mathfrak{S}}(\mathfrak{M}) \hookrightarrow \mathrm{Hom}_{\mathfrak{O},\varphi,\mathrm{Fil}}(\varphi^*\mathcal{M}, B^+_{\mathrm{cris},F}) \hookrightarrow \mathrm{Hom}_{\mathfrak{O}_\alpha,\varphi,\mathrm{Fil}}(\mathfrak{O}_\alpha \otimes_{\mathfrak{O}} \varphi^*\mathcal{M}, B^+_{\mathrm{cris},F}) \\ \simeq \mathrm{Hom}_{\mathfrak{O}_\alpha,\varphi,\mathrm{Fil}}(\mathfrak{O}_\alpha \otimes_{F_0} D, B^+_{\mathrm{cris},F}) \simeq V_{\mathrm{cris},F}(D),$$

where instead of using Lemma 4.2.3, we must appeal to Lemma 4.5.6 and Lemma 5.2.1 (note that *a priori* we know neither that $\mathcal{M}(D) \simeq \mathcal{M}$ nor that D is admissible). Since $\dim_{F_0}(D) = \mathrm{rank}_{\mathfrak{S}_F}\mathfrak{M}$ and ι is injective, we conclude that D is admissible. In particular, $V_{\mathrm{cris},F}(D)$ is crystalline with Hodge-Tate weights in $\{0,1\}$.

It remains to show that ι is compatible with the given actions of $G = G_K$. By construction, the G-action on $T_{\mathfrak{S}}(\mathfrak{M})$ is induced from the identification

$$T_{\mathfrak{S}}(\mathfrak{M}) \simeq \mathrm{Hom}_{W(R)_F,\varphi}(W(R)_F \otimes_{\mathfrak{S}_F} \mathfrak{M}, W(R)_F) \tag{5.2.1}$$

of (5.1.4) with G-action on the right side that of Lemma 5.1.1. Now the right side of (5.2.1) is clearly naturally isomorphic as an $\mathcal{O}_F[G]$-module to

$$\mathrm{Hom}_{W(R)_F,\varphi,\mathrm{Fil}}(W(R)_F \otimes_{\mathfrak{S}_F} \varphi^*\mathfrak{M}, W(R)_F),$$

which is an $\mathcal{O}_F$-lattice in $\mathrm{Hom}_{B^+_{\mathrm{cris},F},\varphi,\mathrm{Fil}}(B^+_{\mathrm{cris},F} \otimes_{F_0} D, B^+_{\mathrm{cris},F})$. Thus, to prove that ι is G-equivariant, we must show that the G-action on $B^+_{\mathrm{cris},F} \otimes_{\mathfrak{S}_F} \varphi^*\mathfrak{M}$ deduced from Lemma 5.1.1 agrees with the G-action on $B^+_{\mathrm{cris},F} \otimes_{F_0} D$ via the map

$$B^+_{\mathrm{cris},F} \otimes_{F_0} D \xrightarrow{\simeq} B^+_{\mathrm{cris},F} \otimes_{\mathfrak{S}_F} \varphi^*\mathfrak{M}$$

deduced from (4.5.3) (which is an isomorphism thanks to Lemma 4.5.6); here, G acts trivially on D. Equivalently, we must show that the G-action on $B^+_{\mathrm{cris},F} \otimes_{\mathfrak{S}_F} \varphi^*\mathfrak{M}$ provided by Lemma 5.1.1 restricts to the *trivial* action on $D(\mathfrak{M})$, viewed as a subspace of this tensor product again via (4.5.3).

As in the proofs of Lemma 4.5.6 and Lemma 5.1.1, let $\mathfrak{e}_1, \ldots, \mathfrak{e}_d$ be an $\mathfrak{S}_F$-basis of $\mathfrak{M}$ and put $\{e_i := 1 \otimes \mathfrak{e}_i\}$, which is then an $\mathfrak{S}_F$-basis of $\varphi^*\mathfrak{M}$. The proof of Lemma 4.5.6 shows that $(f_1, \ldots, f_d) := (e_1, \ldots, e_d)Y$ is a basis of $D(\mathfrak{M})$ for

$$Y = \lim_{n\to\infty} \varphi(A)\cdots\varphi^n(A)\varphi^n(A_0^{-1})\cdots\varphi(A_0^{-1}).$$

Now for any $g \in G$, by the proof of Lemma 5.1.1 we have the equality $g(e_1, \ldots, e_d) = (e_1, \ldots, e_d)\varphi(X_g)$ with

$$\varphi(X_g) = \lim_{n\to\infty} \varphi(A)\cdots\varphi^n(A)\varphi^n(g(A^{-1}))\cdots\varphi(g(A^{-1})).$$

Thus, $g(f_1,\dots,f_d) = (e_1,\dots,e_d)\varphi(X_g)g(Y) = (e_1,\dots,e_d)\lim_{n\to\infty}\varphi(X_n)g(Y_n)$ with

$$\begin{aligned}\varphi(X_n)g(Y_n) &= \big(\varphi(A)\cdots\varphi^n(A)\varphi^n(g(A^{-1}))\cdots\varphi(g(A^{-1}))\big)\\ &\quad\times\big(\varphi(g(A))\cdots\varphi^n(g(A))\varphi^n(g(A_0^{-1}))\cdots\varphi(g(A_0^{-1}))\big)\\ &= Y_n\end{aligned}$$

In other words, $g(f_1,\dots,f_d) = (f_1,\dots,f_d)$, which completes the proof. □

Proof of Theorem 5.0.10. Thanks to Proposition 5.2.2 and Corollary 5.1.2, we have a contravariant functor $T_{\mathfrak{S}} : \mathrm{Mod}^{\varphi,1}_{\mathfrak{S}_F} \to \mathrm{Rep}^{F\text{-cris},1}_{\mathcal{O}_F}(G)$, which it remains to prove is fully faithful and essentially surjective.

For full-faithfulness, suppose given a map $h : T_{\mathfrak{S}}(\mathfrak{M}) \to T_{\mathfrak{S}}(\mathfrak{M}')$ of $\mathcal{O}_F[G]$-modules. Restricting to $G_{\underline{\pi}}$ gives a map $h\big|_{G_{\underline{\pi}}} : T_{\mathfrak{S}}(\mathfrak{M})\big|_{G_{\underline{\pi}}} \to T_{\mathfrak{S}_F}(\mathfrak{M}')\big|_{G_{\underline{\pi}}}$, and by Corollary 3.2.3 we obtain a morphism $\mathfrak{f} : \mathcal{O}_{\mathcal{E}} \otimes_{\mathfrak{S}_F} \mathfrak{M}' \to \mathcal{O}_{\mathcal{E}} \otimes_{\mathfrak{S}_F} \mathfrak{M}$ with $T_{\mathfrak{S}}(\mathfrak{f}) = h|_{G_{\underline{\pi}}}$. It then suffices to show that $\mathfrak{f}(\mathfrak{M}') \subset \mathfrak{M}$. Arguing as in the proof of Proposition 3.3.5, it suffices to check that if $\mathfrak{M} \subset \mathfrak{M}' \subset \mathcal{O}_{\mathcal{E}}\otimes_{\mathfrak{S}_F}\mathfrak{M}$ then $\mathfrak{M} = \mathfrak{M}'$. Applying $\wedge^d$, we then easily reduce to proving that $\wedge^d\mathfrak{M}$ and $\wedge^d\mathfrak{M}'$ have the same minimal height. By our reductions are now in the situation that $T_{\mathfrak{S}}(\mathfrak{M}) \simeq T_{\mathfrak{S}}(\mathfrak{M}')$ as $\mathcal{O}_F[G]$-modules thanks to Corollary 5.1.2, so by Proposition 5.2.2 we have $D(\mathfrak{M}) \simeq D(\mathfrak{M}')$ as filtered φ-modules. In particular, $\mathrm{Fil}^1 D(\mathfrak{M})_{F_0,K} \simeq \mathrm{Fil}^1 D(\mathfrak{M}')_{F_0,K}$ and the minimal heights of $\wedge^d\mathfrak{M}$ and $\wedge^d\mathfrak{M}'$ are the same by Lemma 5.2.1 (2). Thus, $T_{\mathfrak{S}}$ is fully faithful.

We now show that $T_{\mathfrak{S}}$ is essentially surjective. Fix $T \in \mathrm{Rep}^{F\text{-cris},1}_{\mathcal{O}_F}(G)$, put $V := F\otimes_{\mathcal{O}_F} T$ and let $D := D_{\mathrm{cris},F}(V)$ be the corresponding filtered φ-module. By Corollary 4.5.2, there exists $\mathfrak{M} \in \mathrm{Mod}^{\varphi,1}_{\mathfrak{S}_F}$ with $\mathcal{M}(D) \simeq \mathfrak{O}\otimes_{\mathfrak{S}_F}\mathfrak{M}$ and $\iota : T_{\mathfrak{S}}(\mathfrak{M}) \xrightarrow{\simeq} T|_{G_{\underline{\pi}}}$. It suffices to show that ι is compatible with the actions of G on source and target, with the G-action on the source provided by Corollary 5.1.2. Using Proposition 5.2.2, we obtain an isomorphism of $F[G]$-modules $\iota' : V_{\mathfrak{S}}(\mathfrak{M}) \simeq V_{\mathrm{cris},F}(D(\mathfrak{M}))$, which one verifies is compatible with the identification ι. It therefore remains to check that $D(\mathfrak{M}) \simeq D$ as filtered φ-modules. Thanks to Lemma 4.2.2 and Lemma 4.5.6, we can identify each of D and $D(\mathfrak{M})$ as the image of the unique φ-equivariant section to projection $\varphi^*\mathcal{M}(D) \twoheadrightarrow \varphi^*\mathcal{M}(D)/u\varphi^*\mathcal{M}(D)$, which gives $D \simeq D(\mathfrak{M})$ as φ-modules. Thus, it remains to prove that $\mathrm{Fil}^i D_K = \mathrm{Fil}^i D(\mathfrak{M})_K$ for all $i > 0$, or equivalently that $\mathrm{Fil}^1 D_{F_0,K} = \mathrm{Fil}^1 D(\mathfrak{M})_{F_0,K}$. Thanks to Corollary 4.2.4, the projection $\psi_\pi : \varphi^*\mathcal{M}(D) \twoheadrightarrow \varphi^*\mathcal{M}(D)/E\varphi^*\mathcal{M}(D) \simeq D_{F_0,K}$ is compatible with filtrations, and one checks using the very definition of $\mathrm{Fil}^1\varphi^*\mathcal{M}(D)$ that $x \in \mathrm{Fil}^1\varphi^*\mathcal{M}(D)$ if and only if $\psi_\pi(x) \in \mathrm{Fil}^1 D_{F_0,K}$. Thus, $\mathrm{Fil}^1 D_{F_0,K} = \psi_\pi(\mathrm{Fil}^1\varphi^*\mathcal{M}(D)) = \mathrm{Fil}^1 D(\mathfrak{M})_{F_0,K}$, as desired. □

Remark 5.2.3. In the classical situation, let S be the p-adic completion of the divided-power envelope of the surjection $W(k)[\![u]\!] \twoheadrightarrow \mathcal{O}_K$ sending u to π. If $\mathfrak{M}$ is the Kisin module attached to a Barsotti–Tate group H over $\mathcal{O}_K$, then one can show ([Kis06, §2.2.3]) that there is a functorial isomorphism of Breuil modules $\varphi^*\mathfrak{M} \otimes_{\mathfrak{S}} S \simeq \mathbb{D}(H)_S$, where $\mathbb{D}(H)$ is the Dieudonné crystal attached to H, which gives a *geometric interpretation* of $\mathfrak{M}$ in terms of the crystalline cohomology of H. It is natural to ask for such an interpretation in the general case, for arbitrary F and $f(u)$ as in the introduction of this paper. If $F/\mathbb{Q}_p$ is unramified, then this interpretation is provided by [CL14]. However, for F ramified over $\mathbb{Q}_p$, things are more subtle as it is necessary to use the $\mathcal{O}$-divided powers of Faltings [Fal02]. For general F, A. Henniges has obtained the analogous relation with the Dieudonné crystal under the restriction $v_F(a_1) = 1$ (the so-called *Lubin–Tate* setting) in his Ph. D thesis. The general case remains open, but we nonetheless conjecture that one has a natural isomorphism $A_{\mathrm{cris},F} \otimes_{\mathfrak{S}_F} \varphi^*\mathfrak{M} \simeq \mathbb{D}(H)_{A_{\mathrm{cris}}}$, and expect to be able to prove this conjecture using the ideas of §6.3.

6. Further Questions

As Theorem 1.0.1 and Theorem 1.0.2 provide the foundations of the theory of Kisin modules and its variants (*e.g.* the theory of $(\varphi, \hat{G})$-modules as in [Liu10]), it is natural to ask to what extent we can extend these theories to accommodate general F and $f(u)$. In this section, we list some questions that are natural next steps to consider in furthering the general theory we have laid out in this paper.

6.1. The case $q = p^s$. Recall the setup of the introduction: $F/\mathbb{Q}_p$ is an arbitrary finite extension with uniformizer ϖ and residue field k_F of cardinality $q = p^s$, and $f(u) \in \mathcal{O}_F[\![u]\!]$ is any power series $f(u) = a_1 u + \cdots$ satisfying $f(u) \equiv u^q \bmod \varpi$. We allow K to be any finite extension of F with uniformizer $\pi = \pi_0$ and residue field $k \supseteq k_F$, and consider the Frobenius-iterate extension $K_{\underline{\pi}}$ formed by adjoining to K a choice of f-compatible system $\{\pi_n\}_n$ in $\overline{K}$ with $f(\pi_n) = \pi_{n-1}$. Such extensions and their associated norm fields are considered in [CD15] and [CDL]. In this paper, we have restricted ourselves to $q = p$, or what is the same, that $F/\mathbb{Q}_p$ is totally ramified. Certainly this restriction is unnecessary, and we are confident that the results of this paper can be adapted to the general case of arbitrary F with minor modifications. In particular, in this general case, for any $W(k)$-algebra A we set $A_F := A \otimes_{W(k_F)} \mathcal{O}_F$, and we equip $\mathfrak{S}_F$ with the "q-power Frobenius" φ_q which acts on F-trivially, acts on $W(k)$ via $\varphi^s_{W(k)}$ and sends u to $f(u)$. We write $F_0 := K_0 F$ and again denote by $E \in \mathcal{O}_{F_0}[u]$ the minimal polynomial of π over F_0. Then our theory should be able to be adapted to functorially associate Kisin modules of finite E-height to $\mathcal{O}_F$-lattices in F-crystalline G-representations. We note that such a theory is already known in the "Lubin–Tate" case that $v_F(a_1) = 1$ and $K \subseteq F_{\underline{\varpi}}$ thanks to the work of Kisin and Ren [KR09], but that there are many details in our general setup that still need to be checked.

6.2. Semi-stable representations and Breuil theory. In the classical situation, Theorem 4.4.1 includes semi-stable representations. This fact is one of the key inputs for Breuil's classification of lattices in semistable representations via strongly divisible lattices over S (see [Liu08]). It is therefore natural to ask if Theorem 4.4.1 remains valid for semi-stable representations and general $f(u)$. This appears to be a rather nontrivial question, as the case of semi-stable representations requires a monodromy operator. But for general F and f, we do not even know how to define a reasonable monodromy operator over $\mathfrak{S}_F$ (*i.e.*, one satisfying $N\varphi = p\varphi N$ as in the classical situation). New ideas are needed for this direction.

6.3. Comparison between different choices of $f(u)$. For a fixed F-crystalline representation V of G and a fixed uniformizer $\pi \in K$, we may select different $f(u)$. It is then natural to ask for the relationship between the associated Kisin modules attached to V and $f(u)$, as f varies. Motivated by [Liu13a], we conjecture that all such Kisin modules become isomorphic after base change to $W(R)_F$. Note that if true, this result provides a proof of the conjecture mentioned in Remark 5.2.3, because we know that $A_{\mathrm{cris},F} \otimes_{\mathfrak{S}_F} \varphi^*\mathfrak{M} \simeq \mathbb{D}(H)(A_{\mathrm{cris}})$ in the classical situation. To prove such comparison results, the key point is to generalize [Liu07, Theorem 3.2.2] to allow general $f(u)$. This is likely relatively straightforward, as we have recovered many results of [Liu07] in §3 already.

6.4. Torsion theory. A major advantage of the theory of Kisin modules is that it provides a powerful set of tools for dealing with torsion representations. It is therefore natural to try and rebuild the torsion theory in our general situation, and we hope that such a theory will have some striking applications, for example, to the computation of the reduction of potentially crystalline representations as discussed in the introduction. One obvious initial goal is to establish the equivalence between torsion Kisin modules of height 1 and finite flat group schemes over $\mathcal{O}_K$; this would be achievable quickly once we know the truth of the conjecture formulated in Remark 5.2.3.

References

[BB10] Laurent Berger and Christophe Breuil, *Sur quelques représentations potentiellement cristallines de* $\mathrm{GL}_2(\mathbf{Q}_p)$, Astérisque (2010), no. 330, 155–211.

[BCDT01] Christophe Breuil, Brian Conrad, Fred Diamond, and Richard Taylor, *On the modularity of elliptic curves over* $\mathbf{Q}$*: wild 3-adic exercises*, J. Amer. Math. Soc. 14 (2001), no. 4, 843–939 (electronic).

[Ber02] Laurent Berger, *Représentations p-adiques et équations différentielles*, Invent. Math. 148 (2002), no. 2, 219–284.

[Ber04] ———, *Limites de représentations cristallines*, Compos. Math. 140 (2004), no. 6, 1473–1498.

[Ber08] ———, *Construction de* (ϕ, Γ)*-modules: représentations p-adiques et B-paires*, Algebra Number Theory 2 (2008), no. 1, 91–120.

[Ber14] ———, *Multivariable* (φ, Γ) *-modules and locally analytic vectors*, Duke Math J., *to appear* (2016).

[Bre98] Christophe Breuil, *Schémas en groupes et corps des normes*, Unpublished, available from http://www.math.u-psud.fr/˜breuil/PUBLICATIONS/groupesnormes.pdf (1998).

[BTR13] Alexander Beilinson and Floric Tavares Ribeiro, *On a theorem of Kisin*, C. R. Math. Acad. Sci. Paris 351 (2013), no. 13-14, 505–506.

[Car] Xavier Caruso, *Représentations galoisiennes p-adiques et (φ,τ)-modules*, Duke Math. J. 162 (2013), no. 13, 2525–2607.

[CD15] Bryden Cais and Christopher Davis, *Canonical Cohen rings for norm fields*, Int. Math. Res. Not. IMRN (2015), no. 14, 5473–5517.

[CDL] Bryden Cais, Christopher Davis, and Jonathan Lubin, *A characterization of strictly APF extensions*, J. Théorie des Nombres de Bordeaux, *to appear* (2016).

[CL14] Bryden Cais and Eike Lau, *Dieudonne crystals and Wach modules for p-divisible groups*, arXiv preprint arXiv:1412.3174 (2014).

[CI99] Robert Coleman and Adrian Iovita, *The Frobenius and monodromy operators for curves and abelian varieties*, Duke Math. J. 97 (1999), no. 1, 171–215.

[Col99] Pierre Colmez, *Représentations cristallines et représentations de hauteur finie*, J. Reine Angew. Math. 514 (1999), 119–143.

[Col02] ———, *Espaces de Banach de dimension finie*, J. Inst. Math. Jussieu 1 (2002), no. 3, 331–439.

[Fal02] Gerd Faltings, *Group schemes with strict $\mathscr{O}$-action*, Mosc. Math. J. 2 (2002), no. 2, 249–279.

[Fon90] Jean-Marc Fontaine, *Représentations p-adiques des corps locaux. I*, The Grothendieck Festschrift, Vol. II, Progr. Math., vol. 87, Birkhäuser Boston, Boston, MA, 1990, pp. 249–309.

[Fon94a] ———, *Le corps des périodes p-adiques*, Astérisque (1994), no. 223, 59–111.

[Fon94b] ———, *Représentations p-adiques semi-stables*, Astérisque (1994), no. 223, 113–184, With an appendix by Pierre Colmez, Périodes p-adiques (Bures-sur-Yvette, 1988).

[GLS14] Toby Gee, Tong Liu, and David Savitt, *The Buzzard-Diamond-Jarvis conjecture for unitary groups*, J. Amer. Math. Soc. 27 (2014), no. 2, 389–435.

[Ked04] Kiran S. Kedlaya, *A p-adic local monodromy theorem*, Ann. of Math. (2) 160 (2004), no. 1, 93–184.

[Kis06] Mark Kisin, *Crystalline representations and F-crystals*, Algebraic geometry and number theory, Progr. Math., vol. 253, Birkhäuser Boston, Boston, MA, 2006, pp. 459–496.

[Kis08] ———, *Potentially semi-stable deformation rings*, J. Amer. Math. Soc. 21 (2008), no. 2, 513–546.

[Kis09a] ———, *The Fontaine-Mazur conjecture for* GL_2, J. Amer. Math. Soc. 22 (2009), no. 3, 641–690.

[Kis09b] ———, *Moduli of finite flat group schemes, and modularity*, Ann. of Math. (2) 170 (2009), no. 3, 1085–1180.

[KR09] Mark Kisin and Wei Ren, *Galois representations and Lubin-Tate groups*, Doc. Math. 14 (2009), 441–461.

[Laz62] Michel Lazard, *Les zéros des fonctions analytiques d'une variable sur un corps valué complet*, Inst. Hautes Études Sci. Publ. Math. (1962), no. 14, 47–75.

[Liu07] Tong Liu, *Torsion p-adic Galois representations and a conjecture of Fontaine*, Ann. Sci. École Norm. Sup. (4) 40 (2007), no. 4, 633–674.

[Liu08] ———, *On lattices in semi-stable representations: a proof of a conjecture of Breuil*, Compos. Math. 144 (2008), no. 1, 61–88.

[Liu10] ———, *A note on lattices in semi-stable representations*, Mathematische Annalen 346 (2010), no. 1, 117–138.

[Liu11] ———, *Filtration assocaited to torsion semi-stable representations*, Annales de l'Institut Fourier, *to appear* (2011).

[Liu13a] ———, *Compatibility of Kisin modules for different uniformizers*, J. Reine Angew. Math., *to appear* (2013).

[Liu13b] ———, *The correspondence between Barsotti-Tate groups and Kisin modules when* $p = 2$, J. Théor. Nombres Bordeaux 25 (2013), no. 3, 661–676.

[Wac96] Nathalie Wach, *Représentations p-adiques potentiellement cristallines*, Bull. Soc. Math. France 124 (1996), no. 3, 375–400.

[Win83] Jean-Pierre Wintenberger, *Le corps des normes de certaines extensions infinies de corps locaux; applications*, Ann. Sci. École Norm. Sup. (4) 16 (1983), no. 1, 59–89.

Bryden Cais
Department of Mathematics
University of Arizona
Tucson, Arizona 85721
USA
cais@math.arizona.edu

Tong Liu
Department of Mathematics
Purdue University
West Lafayette IN 47907
USA
tongliu@math.purdue.edu

Equivariant Poincaré Series and Topology of Valuations

A. Campillo, F. Delgado[1], and S.M. Gusein-Zade [2]

Received: September 7, 2015
Revised: February 10, 2016

Communicated by Thomas Peternell

Abstract. The equivariant with respect to a finite group action Poincaré series of a collection of r valuations was defined earlier as a power series in r variables with the coefficients from a modification of the Burnside ring of the group. Here we show that (modulo simple exceptions) the equivariant Poincaré series determines the equivariant topology of the collection of valuations.

2010 Mathematics Subject Classification: 14B05, 13A18, 14R20, 16W70.
Keywords and Phrases: finite group actions, Poincaré series, plane valuations, equivariant topology

1 Introduction

A definition of the Poincaré series of a multi-index filtration was first given in [3] (for filtrations defined by collections of valuations). It is a formal power series in several variables with integer coefficients, i.e., an element of the ring $\mathbb{Z}[[t_1, \ldots, t_r]]$. In [1] it was shown that, for the filtration defined by the curve valuations corresponding to the irreducible components of a plane curve singularity, the Poincaré series coincides with the Alexander polynomial in several variables of the corresponding algebraic link: the intersection of the curve with a small sphere in $\mathbb{C}^2$ centred at the origin. This relation was obtained by a direct computation of the both sides in the same terms. Up to now there exist

[1]Supported by the grants MTM2012-36917-C03-01 / 02 (first grant with the help of FEDER Program).
[2]Supported by the grants RFBR–13-01-00755, NSh–5138.2014.1.

no conceptual proof of it. The Alexander polynomial in several variables of an algebraic link (and therefore the Poincaré series of the corresponding collection of valuations) determines the topological type of the corresponding plane curve singularity. In [2] the definition of the Poincaré series was reformulated in terms of an integral with respect to the Euler characteristics (over an infinite dimensional space).

The desire to understand deeper this relation led to attempts to find an equivariant version of it (for actions of a finite group G) and thus to define equivariant versions of the Poincaré series and of the Alexander polynomial. Some equivariant versions of the monodromy zeta-function (that is of the Alexander polynomial in one variable) were defined in [9] and [10]. Equivariant versions of the Poincaré series were defined in [4], [5] and [7].

In some constructions of equivariant analogues of invariants (especially those related to the Euler characteristic) the role of the ring of integers $\mathbb{Z}$ (where the Euler characteristic takes values) is played by the Burnside ring $A(G)$ of the group G. Therefore it would be attractive to define equivariant versions of the Poincaré series as elements of the ring $A(G)[[t_1, \ldots, t_r]]$ (or of a similar one). The equivariant versions of the monodromy zeta functions defined in [9] and [10] are formal power series with the coefficients from $A(G) \otimes \mathbb{Q}$ and $A(G)$ respectively.

In [4] the equivariant Poincaré series was defined as an element of the ring $R_1(G)[[t_1, \ldots, t_r]]$ of formal power series in $t_1, \ldots, t_r$ with the coefficients from the subring $R_1(G)$ of the ring $R(G)$ of complex representations of the group G generated by the one-dimensional representations. This Poincaré series turned out to be useful for some problems: see, e.g., [8], [11]. However, it seems to be rather "degenerate", especially for non-abelian groups.

In [5] the G-equivariant Poincaré series $P^G_{\{\nu_i\}}$ of a collection of valuations (or order functions) $\{\nu_i\}$ was not in fact a series, but an element of the Grothendieck ring of so called locally finite (G, r)-sets. This Grothendieck ring was rather big and complicated, the Poincaré series $P^G_{\{\nu_i\}}$ was rather complicated as well and contained a lot of information about the valuations and the G-action. In particular, for curve and divisorial valuations on the ring $\mathcal{O}_{\mathbb{C}^2,0}$ of functions in two variables the information contained in this Poincaré series was (almost) sufficient to restore the action of G on $\mathbb{C}^2$ and the G-equivariant topology of the set of valuations: [6].

In [7] the equivariant Poincaré series $P^G_{\{\nu_i\}}(t_1, \ldots, t_r)$ was defined as an element of the ring $\widetilde{A}(G)[[t_1, \ldots, t_r]]$ of formal power series in the variables $t_1, \ldots, t_r$ with the coefficients from a certain modification $\widetilde{A}(G)$ of the Burnside ring $A(G)$ of the group G. A simple reduction of this Poincaré series is an element of the ring $A(G)[[t_1, \ldots, t_r]]$. Thus it is somewhat close to the ("idealistic") model discussed above. However, in order to define the equivariant Poincaré series of this form, it was necessary to lose quite a lot of information about the individual valuations from the collection. (It is possible to say that one used averaging of the information over the group.) Thus it was not clear how much

information does it keep.
Equivariant topology of germs of plane curves seems to be much more involved than the usual (non-equivariant) one. For example, it is unclear whether the equivariant topology of a collection of curves always determines the equivariant Poincaré series of the collection. Here we discuss to which extent the G-equivariant Poincaré series from [7] determines the topology of a set of plane valuations. We show that the G-equivariant Poincaré series of a collection of divisorial valuations determines the equivariant topology of this collection (in a natural "weak" sense: see below). We also show that with some minor exceptions the equivariant Poincaré series of a collection of curve valuations determines the weak equivariant topology of the collection. (This answer resembles the one from [6]. However reasons for that (and thus the proofs) are quite different. The version of the equivariant Poincaré series considered in [6] is apriori a much more fine invariant than that considered here.)
The G-equivariant Poincaré series $P^G_{\{\nu_i\}}$ considered in [5] depends essentially on the set of valuations defining the filtration. In particular, the substitution of one of them (say, ν_i) by its shift $a^*\nu_i$, $a \in G$, changes the G-equivariant Poincaré series $P^G_{\{\nu_i\}}$. The Poincaré series $P^G_{\{\nu_i\}}(\underline{t})$ considered in [7] depends not on the valuations ν_i themselves, but on their G-orbits. The substitution of one of them by its shift does not change the G-equivariant Poincaré series $P^G_{\{\nu_i\}}(\underline{t})$. Therefore this series cannot determine the G-topology of a collection of divisorial and/or of curve valuations on $\mathcal{O}_{\mathbb{C}^2,0}$ in the form defined in [6]. One has to modify this notion a little bit.
Assume first that we consider sets of curve valuations. Let $\{C_i\}_{i=1}^r$ and $\{C'_i\}_{i=1}^r$ be two collections of branches (that is of irreducible plane curve singularities) in the complex plane $(\mathbb{C}^2, 0)$ with an action of a finite group G. We shall say that these collections are *weakly G-topological equivalent* if there exists a G-invariant germ of a homeomorphism $\psi : (\mathbb{C}^2, 0) \to (\mathbb{C}^2, 0)$ such that for each $i = 1, \ldots, r$ one has $\psi(C_i) = a_i C'_i$ with an element $a_i \in G$ (i.e if the image of the G-orbit of the branch C_i coincides with the G-orbit of the branch C'_i). To formulate an analogue of this definition for collections of divisorial valuations, one can describe a divisorial valuation ν on $\mathcal{O}_{\mathbb{C}^2,0}$ by a pair of curvettes intersecting the corresponding divisor (transversally) at different points. The corresponding pair of curvettes allows to determine the divisor as the last one (and so the unique one with self-intersection equal to -1) appearing in the minimal embedded resolution of them. Two collections of divisorial valuations $\{\nu_i\}_{i=1}^r$ and $\{\nu'_i\}_{i=1}^r$ described by the corresponding collections of curvettes $\{L_{ij}\}_{i=1,j=1,2}^r$ and $\{L'_{ij}\}_{i=1,j=1,2}^r$ respectively are weakly G-topologically equivalent if there exists a G-invariant germ of a homeomorphism $\psi : (\mathbb{C}^2, 0) \to (\mathbb{C}^2, 0)$ such that for each $i = 1, \ldots, r$ one has $\psi(L_{ij}) = a_i L'_{ij}$ for $j = 1, 2$ and an element $a_i \in G$.
One has an obvious analogue of Theorem 2.9 from [6]. This means that, for a fixed representation of the group G on $\mathbb{C}^2$, the weak topology of a collection of curve or/and divisorial valuations on $\mathcal{O}_{\mathbb{C}^2,0}$ is determined by the G-resolution graph Γ^G of the collection (where not individual branches or/and divisors, but their orbits are indicated) plus the correspondence between the tails of this

graph emerging from special points of the first component of the exceptional divisor with these special points (see below).

2 Equivariant Poincaré series

Let us briefly recall the definition of the G-equivariant Poincaré series $P^G_{\{\nu_i\}}(t_1, \ldots, t_r)$ of a collection of order functions on the ring $\mathcal{O}_{V,0}$ of germs of functions on $(V, 0)$ and the equation for it in terms of a G-equivariant resolution of curve or/and divisorial plane valuations which will be used here.

Definition: A finite *equipped G-set* is a pair $\widetilde{X} = (X, \alpha)$ where:

- X is a finite G-set;
- α associates to each point $x \in X$ a one-dimensional representation α_x of the isotropy subgroup $G_x = \{a \in G : ax = x\}$ of the point x so that, for $a \in G$, one has $\alpha_{ax}(b) = \alpha_x(a^{-1}ba)$, where $b \in G_{ax} = aG_xa^{-1}$.

Let $\widetilde{A}(G)$ be the Grothendieck group of finite equipped G-sets. The cartesian product defines a ring structure on it. The class of an equipped G-set $\widetilde{X}$ in the Grothendieck ring $\widetilde{A}(G)$ will be denoted by $[\widetilde{X}]$. As an abelian group $\widetilde{A}(G)$ is freely generated by the classes of the irreducible equipped G-sets $[G/H]_\alpha$ for all the conjugacy classes $[H]$ of subgroups of G and for all one-dimensional representations α of H (a representative of the conjugacy class $[H] \in \mathrm{Conjsub}\, G$). There is a natural homomorphism ρ from the ring $\widetilde{A}(G)$ to the Burnside rings $A(G)$ of the group G defined by forgetting the one-dimensional representation corresponding to the points. The reduction $\hat{\rho} : \widetilde{A}(G) \to \mathbb{Z}$ is defined by forgetting the representations and the G-action. There are natural pre-λ-structure on a rings $A(G)$ and $\widetilde{A}(G)$ which give sense for the expressions of the form $(1-t)^{-[X]}$, $[X] \in A(G)$, and $(1-t)^{-[\widetilde{X}]}$, $[\widetilde{X}] \in \widetilde{A}(G)$ respectively: see [7]. Both ρ and $\hat{\rho}$ are homomorphisms of pre-λ-rings.
Let $(V, 0)$ be a germ of a complex analytic space with an action of a finite group G and let $\mathcal{O}_{V,0}$ be the ring of germs of functions on it. Without loss of generality we assume that the G-action on $(V, 0)$ is faithful. The group G acts on $\mathcal{O}_{V,0}$ by $a^*f(z) = f(a^{-1}z)$ ($z \in V$, $a \in G$). A valuation ν on the ring $\mathcal{O}_{V,0}$ is a function $\nu : \mathcal{O}_{V,0} \to \mathbb{Z}_{\geq 0} \cup \{+\infty\}$ such that:

1) $\nu(\lambda f) = \nu(f)$ for $\lambda \in \mathbb{C}^*$;
2) $\nu(f + g) \geq \min\{\nu(f), \nu(g)\}$;
3) $\nu(fg) = \nu(f) + \nu(g)$.

A function $\nu : \mathcal{O}_{V,0} \to \mathbb{Z}_{\geq 0} \cup \{+\infty\}$ which possesses the properties 1) and 2) is called an *order function*.
Let $\nu_1, \ldots, \nu_r$ be a collection of order functions on $\mathcal{O}_{V,0}$. It defines an r-index filtration on $\mathcal{O}_{V,0}$:

$$J(\underline{v}) = \{h \in \mathcal{O}_{V,0} : \underline{\nu}(h) \geq \underline{v}\} ,$$

where $\underline{v} = (v_1, \ldots, v_r) \in \mathbb{Z}^r_{\geq 0}$, $\underline{\nu}(h) = (\nu_1(h), \ldots, \nu_r(h))$ and $\underline{v}' = (v'_1, \ldots, v'_r) \geq \underline{v}'' = (v''_1, \ldots, v''_r)$ if and only if $v'_i \geq v''_i$ for all i.

Let $\omega_i : \mathcal{O}_{V,0} \to \mathbb{Z}_{\geq 0} \cup \{+\infty\}$ be defined by $\omega_i = \sum_{a \in G} a^*\nu_i$. The functions ω_i are G-invariant (they are not, in general, order functions). For an element $h \in \mathbb{P}\mathcal{O}_{V,0}$, that is for a function germ considered up to a constant factor, let G_h be the isotropy subgroup $G_h = \{a \in G : a^*h = \alpha_h(a)h\}$ and let $Gh \cong G/G_h$ be the orbit of h in $\mathbb{P}\mathcal{O}_{V,0}$. The correspondence $a \mapsto \alpha_h(a) \in \mathbb{C}^*$ determines a one-dimensional representation α_h of the subgroup G_h. Let $\widetilde{X}_h = [G/G_h]_{\alpha_h}$ be the element of the ring $\widetilde{A}(G)$ represented by the G-set Gh with the representation α_{a^*h} associated to the point $a^*h \in Gh$ $(a \in G)$. The correspondence $h \mapsto \widetilde{X}_h$ defines a function $(\widetilde{X})$ on $\mathbb{P}\mathcal{O}_{V,0}/G$ with values in $\widetilde{A}(G)$. The *equivariant Poincaré series* $P^G_{\{\nu_i\}}(\underline{t})$ of the collection $\{\nu_i\}$ is defined by the equation

$$P^G_{\{\nu_i\}}(\underline{t}) = \int_{\mathbb{P}\mathcal{O}_{V,0}/G} \widetilde{X}_h \underline{t}^{\underline{\omega}(h)} d\chi \in \widetilde{A}(G)[[t_1, \ldots, t_r]] \,, \tag{1}$$

where $\underline{t} := (t_1, \ldots, t_r)$, $\underline{t}^{\underline{\omega}(h)} = t_1^{\omega_1(h)} \cdot \ldots \cdot t_r^{\omega_r(h)}$, $t_i^{+\infty}$ should be regarded as 0. The precise meaning of this integral see in [7].

Applying the reduction homomorphism $\rho : \widetilde{A}(G) \to A(G)$ to the Poincaré series $P^G_{\{\nu_i\}}(\underline{t})$, i.e. to its coefficients, one gets the series $\rho P^G_{\{\nu_i\}}(\underline{t}) \in A(G)[[t_1, \ldots, t_r]]$, i.e. a power series with the coefficients from the (usual) Burnside ring. Applying the homomorphism $\widehat{\rho} : \widetilde{A}(G) \to \mathbb{Z}$ one gets the series $\widehat{\rho} P^G_{\{\nu_i\}}(\underline{t}) \in \mathbb{Z}[[t_1, \ldots, t_r]]$. One has

$$\widehat{\rho} P^G_{\{\nu_i\}}(\underline{t}) = P_{\{a^*\nu_i\}}(t_1, \ldots, t_1, t_2, \ldots, t_2, \ldots, t_r, \ldots, t_r) \,,$$

where $P_{\{a^*\nu_i\}}(\bullet)$ is the usual (non-equivariant) Poincaré series of the collection of $|G|r$ order functions $\{a^*\nu_1, a^*\nu_2, \ldots, a^*\nu_r | a \in G\}$ (each group of equal variables in $P_{\{a^*\nu_i\}}$ consists of $|G|$ of them).

Now assume that a finite group G acts linearly on $(\mathbb{C}^2, 0)$ and let ν_i, $i = 1, \ldots, r$, be either a curve or a divisorial valuation on $\mathcal{O}_{\mathbb{C}^2,0}$. We shall write $I_0 = \{1, 2, \ldots, r\} = I' \sqcup I''$, where $i \in I'$ if and only if the corresponding valuation ν_i is a curve one. For $i \in I'$, let $(C_i, 0)$ be the plane curve defining the valuation ν_i.

A *G-equivariant resolution* (or a *G-resolution* for short) of the collection $\{\nu_i\}$ of valuations is a proper complex analytic map $\pi : (\mathcal{X}, \mathcal{D}) \to (\mathbb{C}^2, 0)$ from a smooth surface $\mathcal{X}$ with a G-action such that:

1) π is an isomorphism outside of the origin in $\mathbb{C}^2$;

2) π commutes with the G-actions on $\mathcal{X}$ and on $\mathbb{C}^2$;

3) the total transform $\pi^{-1}(\bigcup\limits_{i \in I',\, a \in G} aC_i)$ of the curve $GC = G(\bigcup\limits_{i \in I'} C_i)$ is a normal crossing divisor on $\mathcal{X}$ (in particular, the exceptional divisor $\mathcal{D} = \pi^{-1}(0)$ is a normal crossing divisor as well);

4) for each branch C_i, $i \in I'$, its strict transform $\widetilde{C}_i$ is a germ of a smooth curve transversal to the exceptional divisor $\mathcal{D}$ at a smooth point x of it and is invariant with respect to the isotropy subgroup $G_x = \{g \in G : gx = x\}$ of the point x;

5) for each $i \in I''$, the exceptional divisor $\mathcal{D} = \pi^{-1}(0)$ contains the divisor defining the divisorial valuation ν_i.

A G-resolution can be obtained by a G-invariant sequence of blow-ups of points. The action of the group G on the first component of the exceptional divisor can either be trivial (this may happen only if G is cyclic) or have fixed points of (proper) subgroups of G. (If G is abelian, these are the fixed points of G itself.) These points are called *special.*

Let $\overset{\circ}{\mathcal{D}}$ be the "smooth part" of the exceptional divisor $\mathcal{D}$ in the total transform $\pi^{-1}(GC)$ of the curve GC, i.e., $\mathcal{D}$ itself minus all the intersection points of its components and all the intersection points with the components of the strict transform of the curve GC. For $x \in \overset{\circ}{\mathcal{D}}$, let $\widetilde{L}_x$ be a germ of a smooth curve on $\mathcal{X}$ transversal to $\overset{\circ}{\mathcal{D}}$ at the point x and invariant with respect to the isotropy subgroup G_x of the point x. The image $L_x = \pi(\widetilde{L}_x) \subset (\mathbb{C}^2, 0)$ is called a *curvette* at the point x. Let the curvette L_x be given by an equation $h_x = 0$, $h_x \in \mathcal{O}_{\mathbb{C}^2,0}$. Without loss of generality one can assume that the function germ h_x is G_x-equivariant. Moreover we shall assume that the germs h_x associated to different points $x \in \overset{\circ}{\mathcal{D}}$ are choosen so that $h_{ax}(a^{-1}z)/h_x(z)$ is a constant (depending on a and x).

Let E_σ, $\sigma \in \Gamma$, be the set of all irreducible components of the exceptional divisor $\mathcal{D}$ (Γ is a G-set itself). For σ and δ from Γ, let $m_{\sigma\delta} := \nu_\sigma(h_x)$, where ν_σ is the corresponding divisorial valuation, h_x is the germ defining the curvette at a point $x \in E_\delta \cap \overset{\circ}{\mathcal{D}}$. One can show that the matrix $(m_{\sigma\delta})$ is minus the inverse matrix to the intersection matrix $(E_\sigma \circ E_\delta)$ of the irreducible components of the exceptional divisor $\mathcal{D}$. For $i = 1, \ldots, r$, let $m_{\sigma i} := m_{\sigma\delta}$, where E_δ is the component of $\mathcal{D}$ corresponding to the valuation ν_i, i.e. either the component defining the valuation ν_i if ν_i is a divisorial valuation (i.e. if $i \in I''$), or the component intersecting the strict transform of the corresponding irreducible curve C_i if ν_i is a curve valuation (i.e. if $i \in I'$). Let $\underline{m}_\sigma := (m_{\sigma 1}, \ldots, m_{\sigma r}) \in \mathbb{Z}^r_{\geq 0}$, $M_{\sigma i} := \sum_{a \in G} m_{(a\sigma)i}$, $\underline{M}_\sigma := (M_{\sigma 1}, \ldots, M_{\sigma r}) = \sum_{a \in G} \underline{m}_{a\sigma}$.

Let $\widehat{\mathcal{D}}$ be the quotient $\overset{\circ}{\mathcal{D}}/G$ and let $p : \overset{\circ}{\mathcal{D}} \to \widehat{\mathcal{D}}$ be the factorization map. Let $\{\Xi\}$ be a stratification of the smooth curve $\widehat{\mathcal{D}}$ such that:

1) each stratum Ξ is connected;

2) for each point $\widehat{x} \in \Xi$ and for each point x from its pre-image $p^{-1}(\widehat{x})$, the conjugacy class of the isotropy subgroup G_x of the point x is the same, i.e., depends only on the stratum Ξ.

The condition 2) is equivalent to say that the factorization map $p : \overset{\circ}{\mathcal{D}} \to \widehat{\mathcal{D}}$ is a (non-ramified) covering over each stratum Ξ. The condition 1) implies that the inverse image in $\overset{\circ}{\mathcal{D}}$ of each stratum Ξ lies in the orbit of one component E_σ of the exceptional divisor. The element $\underline{M}_\sigma \in \mathbb{Z}^r_{\geq 0}$ depends only on the stratum Ξ and will be denoted by $\underline{M}_\Xi$.

For a point $x \in \overset{\circ}{\mathcal{D}}$, let $\widetilde{X}_x = [G/G_x]_{\alpha_{h_x}} \in \widetilde{A}(G)$. The equipped G-set $\widetilde{X}_x$ is one and the same for all points x from the preimage of a stratum Ξ and therefore it defines an element of $\widetilde{A}(G)$ which we shall denote by $[G/G_\Xi]_{\alpha_\Xi}$. In [7, Theorem 1] it was shown that

$$P^G_{\{\nu_i\}}(\underline{t}) = \prod_{\Xi} \left(1 - \underline{t}^{\underline{M}_\Xi}\right)^{-\chi(\Xi)[G/G_\Xi]_{\alpha_\Xi}} . \tag{2}$$

3 Topology of plane valuations

Let the complex plane $(\mathbb{C}^2, 0)$ be endowed by a faithful linear G-action and let $\{\nu_i\}_{i=1}^r$ be a collection of divisorial valuations on $\mathcal{O}_{\mathbb{C}^2,0}$.

Theorem 1 *The G-equivariant Poincaré series $P^G_{\{\nu_i\}}(\underline{t})$ of the collection $\{\nu_i\}$ of divisorial valuations determines the weak G-equivariant topology of this collection.*

Proof. One has to use the following "projection formula". Let $I = \{i_1, \ldots, i_s\}$ be a subset of the set $\{1, \ldots, r\}$ of the indices numbering the valuations. Then one has

$$P^G_{\{\nu_i\}_{i\in I}}(t_{i_1}, \ldots, t_{i_s}) = P^G_{\{\nu_i\}_{i=1}^r}(t_1, \ldots, t_r)_{|t_i=1 \text{ for } i\notin I} ,$$

i.e. the (G-equivariant) Poincaré series for a subcollection of valuations is obtained from the one for the whole collection by substituting t_i by 1 for all i numbering the valuations which do not participate in the subcollection. (This equation is not valid for other types of valuations, say, for curve ones: see the proof of Theorem 2). The projection formula implies, in particular, that the G-equivariant Poincaré series $P^G_{\{\nu_i\}}(\underline{t})$ of a collection of divisorial valuations determines the G-equivariant Poincaré series (in one variable) of each individual valuation from it.

First we shall show that the Poincaré series $P^G_{\{\nu_i\}}(\underline{t})$ determines the G-resolution graph of the collection of valuations. It turns out that the necessary information about the G-equivariant resolution graph can be restored from the ρ-reduction $\rho P^G_\nu(t)$ of the G-equivariant Poincaré series $P^G_\nu(t)$ (i.e. the series from $A(G)[[t]]$ obtained by forgetting the one-dimensional representations associated with the G-orbits). Therefore we shall start with considering it.

First let us prove the statement for one divisorial valuation. The dual graph Γ^G of the minimal G-equivariant resolution of a divisorial valuation ν looks like in Fig. 1. This means the following.

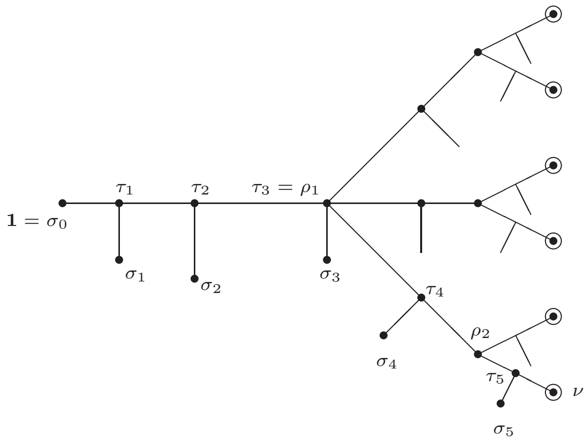


Figure 1: The dual equivariant resolution graph Γ^G of the valuation ν.

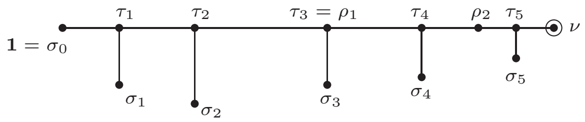


Figure 2: The dual resolution graph Γ of the valuation ν.

The standard (non-equivariant, minimal) dual resolution graph Γ of the valuation ν looks like in Fig. 2. The vertices σ_q, $q = 0, 1, \ldots, g$, are the dead ends of the graph (g is the number of the Puiseux pairs of a curvette corresponding to the valuation, $\sigma_0 = \mathbf{1}$ is the first component of the exceptional divisor), the vertices τ_i, $q = 1, \ldots, g$, are the rupture points, the vertex ν corresponds to the divisorial valuation under consideration. (The vertex ν may coincide with τ_g.) The set of vertices of the graph Γ is ordered according to the order of the birth of the corresponding components of the exceptional divisor. On $[\sigma_0, \nu]$ (the geodesic from $\sigma_0 = \mathbf{1}$ to ν) this order is the natural one: $\delta_1 < \delta_2$ if and only if the vertex δ_1 lies on $[\sigma_0, \delta_2]$.

The integers m_{σ_q}, $q = 0, 1, \ldots, g$, form the minimal set of generators of the semigroup of values of ν and are traditionally denoted by $\overline{\beta}_q$. One also uses the following notations. $e_q := \gcd(\overline{\beta}_0, \overline{\beta}_1, \ldots, \overline{\beta}_q)$,

$$N_q := \frac{e_{q-1}}{e_q} \left(= \frac{m_{\tau_q}}{m_{\sigma_q}} \right).$$

The graph Γ^G of the minimal G-equivariant resolution consists of $|G|$ copies of graph Γ (numbered by the elements of G) glued together. The gluing is defined by a sequence

$$G = H_0 \supset H_1 \supset H_2 \supset \ldots \supset H_k$$

of subgroups of the group G such that all H_i with $i > 0$ are abelian and H_k is the isotropy group of the valuation ν ($\{a \in G : a^*\nu = \nu\}$) and by a sequence by vertices $\rho_1, \ldots, \rho_k$ of the graph Γ such that all of them lie on the geodesic from σ_0 to ν, $\rho_1 < \rho_2 < \ldots < \rho_k$. (Some of the vertices ρ_i may coincide with some of the vertices τ_j; the vertex ρ_1 may coincide with the initial vertex $\sigma_0 = \mathbf{1}$.) The copies of Γ numbered by the elements a_1 and a_2 from G are glued along the part preceeding ρ_ℓ (i.e., by identifying all the vertices smaller or equal to ρ_ℓ) if $a_1 a_2^{-1} \in H_{\ell-1}$. (In particular the initial vertices $\sigma_0 = \mathbf{1}$ of all the copies are identified.) Pay attention that some of the vertices ρ_i can be inbetween the vertices τ_g and ν. For $q = 1, 2, \ldots, g$, let $j(q)$ be defined by the condition $\rho_{j(q)} < \tau_q \leq \rho_{j(q)+1}$.

For $\delta \in \Gamma^G$ (or for the corresponding $\delta \in \Gamma$), let $M_\delta := \sum_{a \in G} m_{a\delta}$. One can easily see that all the integers M_δ, $\delta \in \Gamma$, are different. (One has $M_{\delta_1} = M_{\delta_2}$ for δ_1 and δ_2 from Γ^G if and only if there exists $a \in G$ such that $\delta_2 = a\delta_1$.) One has $M_{\tau_q} = N_q M_{\sigma_q}$.

The series $\rho P^G_\nu(t)$ is given by the equation

$$\begin{aligned}
\rho P^G_\nu(t) \quad &= \quad \prod_{q=0}^{g} \left(1 - t^{M_{\sigma_q}}\right)^{-[G/H_{j(q)}]} \cdot \prod_{q=1}^{g} \left(1 - t^{N_q M_{\sigma_q}}\right)^{[G/H_{j(q)}]} \times \\
&\times \quad \prod_{j=1}^{\ell} \left(1 - t^{M_{\rho_j}}\right)^{[G/H_j]-[G/H_{j-1}]} \cdot \left(1 - t^{M_\nu}\right)^{-[G/H_k]} .
\end{aligned}$$

The fact that all the integers M_δ are different implies that the exponents M_{σ_q}, $q = 1, \ldots, g$, are among those which participate in the decomposition of the series $\rho P^G_\nu(t)$ with negative cardinalities of the multiplicities. (The multiplicity of a binomial $(1 - t^m)^{s_m}$, $s_m \in A(G)$, is s_m. Its cardinality is the (virtual) number of the points of it.) It is possible that the exponents of this sort include also M_ν corresponding to the divisorial valuation itself.

The subgroups $H_1 \supset H_2 \supset \ldots \supset H_k$ are defined by the multiplicities of all the factors in the decomposition of the series $\rho P^G_\nu(t)$ into the product of the binomials.

The vertex $\sigma_0 = \mathbf{1}$ coincides with ρ_1 if and only if the binomial with the smallest exponent in the decomposition of the series $\rho P^G_\nu(t)$ has a non-negative cardinality of the multiplicity. For $\sigma_q \leq \rho_1$ one has $M_{\sigma_q} = |G| m_{\sigma_q}$ and $M_{\rho_1} = |G| m_{\rho_1}$. These equations give all the generators $\overline{\beta}_q$ of the semigroup of values with $\sigma_q \leq \rho_1$ and also m_{ρ_1}.

For $\ell \geq 1$, let $\sigma_{q(\ell)}$ be the minimal dead end greater than ρ_ℓ (i.e. there are the dead ends $\sigma_{q(\ell)}, \ldots, \sigma_{q(\ell+1)-1}$ inbetween ρ_ℓ and $\rho_{\ell+1}$). Let us consider the dead ends σ_q such that $\rho_1 < \sigma_q < \rho_2$. One has

$$M_{\sigma_{q(1)}} = |H_1| m_{\sigma_{q(1)}} + (|G| - |H_1|) m_{\rho_1} = |H_1| m_{\sigma_{q(1)}} + (M_{\rho_1} - |H_1| m_{\rho_1}) .$$

The smallest multiple of the exponent $M_{\sigma_{q(1)}}$ in a binomial participating in the decomposition of the series $\rho P^G_\nu(t)$ is $M_{\tau_{q(1)}} = N_{q(1)}M_{\sigma_{q(1)}}$. Further, for $\rho_1 < \sigma_{q(1)} < \sigma_{q(1)+1} < \sigma_{q(1)+2} < \cdots \sigma_{q(2)-1} < \rho_2$, one has

$$\begin{aligned}
M_{\sigma_{q(1)+1}} &= |H_1|m_{\sigma_{q(1)+1}} + (M_{\rho_1} - |H_1|m_{\rho_1})N_{q(1)}\,,\\
M_{\sigma_{q(1)+2}} &= |H_1|m_{\sigma_{q(1)}+2} + (M_{\rho_1} - |H_1|m_{\rho_1})N_{q(1)}N_{q(1)+1}\,,\\
&\ldots\\
M_{\rho_2} &= |H_1|m_{\rho_2} + (M_{\rho_1} - |H_1|m_{\rho_1})N_{q(1)}N_{q(1)+1}\cdot\ldots\cdot N_{q(2)-1}\,.
\end{aligned}$$

These equations give all the generators $\overline{\beta}_q$ of the semigroup of values with $\sigma_q < \rho_2$ and also m_{ρ_2}.
Assume that we have determined all the exponents m_{σ_q} for $q < q(\ell)$ and also the exponent m_{ρ_ℓ}. Let us consider the dead ends σ_q such that $\rho_\ell < \sigma_q < \rho_{\ell+1}$. One has

$$\begin{aligned}
M_{\sigma_{q(\ell)}} &= |H_\ell|m_{\sigma_{q(\ell)}} + (M_{\rho_\ell} - |H_\ell|m_{\rho_\ell})\,,\\
M_{\sigma_{q(\ell)+1}} &= |H_\ell|m_{\sigma_{q(\ell)+1}} + (M_{\rho_\ell} - |H_\ell|m_{\rho_\ell})N_{q(\ell)}\,,\\
M_{\sigma_{q(\ell)+2}} &= |H_\ell|m_{\sigma_{q(\ell)+2}} + (M_{\rho_\ell} - |H_\ell|m_{\rho_\ell})N_{q(\ell)}N_{q(\ell)+1}\,,\\
&\ldots\\
M_{\rho_{\ell+1}} &= |H_\ell|m_{\rho_\ell} + (M_{\rho_\ell} - |H_\ell|m_{\rho_\ell})N_{q(\ell)}N_{q(\ell)+1}\cdot\ldots\cdot N_{q(\ell+1)-1}\,.
\end{aligned}$$

These equations give all the generators m_{σ_q} of the semigroup of values with $q < q(\ell+1)$ and also $m_{\rho_{\ell+1}}$.
The described procedure recovers m_{σ_q} for all $q \le g$. If, in the binomials of the decomposition of the series $\rho P^G_\nu(t)$, there are no exponents proportional to M_{σ_g}, one has $\nu = \tau_g$ and the resolution graph Γ is determined by the semigroup $\langle\overline{\beta}_0\overline{\beta}_1, \ldots, \overline{\beta}_g\rangle$. Otherwise the described above procedure permits to determine the exponents m_{ρ_j} with $\rho_j \ge \tau_g$ and m_ν. This gives the G-equivariant resolution graph of one divisorial valuation.

Assume that we have a collection $\{\nu_i\}$ of divisorial valuations, $i = 1, 2, \ldots, r$. To restore the equivariant resolution graph Γ^G of the collection from the resolution graphs of each individual valuation ν_i, one has to determine the separation point δ_{ij} between each two valuations ν_i and ν_j (for simplicity let us assume that $i = 1$, $j = 2$). Let

$$\rho P^G_\nu(t_1, t_2, 1, \ldots, 1) = \prod(1 - t_1^{M_1}t_2^{M_2})^{s_{M_1M_2}}\,, \qquad (3)$$

$s_{M_1M_2} \in \mathbb{Z}$, be the decomposition into the product of the binomials. The separation point δ_{12} corresponds to the maximal exponent in the decomposition (3) with

$$\frac{M_{\delta 1}}{M_{\delta 2}} = \frac{M_{\sigma_0 1}}{M_{\sigma_0 2}}\,.$$

This proves that the reduction $\rho P^G_{\{\nu_i\}}(\underline{t}) \in A(G)[[t_1, \ldots, t_r]]$ of the G-equivariant Poincaré series $P^G_{\{\nu_i\}}(\underline{t})$ determines the minimal G-resolution graph of the set $\{\nu_i\}$ of divisorial valuations.

In order to prove that one can also determine the weak G-topology of the collection of valuations, one has to show how is it possible to restore the representation of the group G on $\mathbb{C}^2$ and the correspondence between (some) tails of the (minimal) G-resolution graph and the special points on the first component of the exceptional divisor. For that one should use the non-reduced Poincaré series $P^G_{\{\nu_i\}}(\underline{t}) \in \widetilde{A}(G)[[t_1, \ldots, t_r]]$ itself. (If there are no special points on the first component of the exceptional divisor (this can happen only if G is cyclic), only the representation of G on $\mathbb{C}^2$ has to be determined.) We follow the scheme described in [6].

Let us consider the case of an abelian group G first. If there are no special points on the first component $E_{\mathbf{1}}$ of the exceptional divisor, all points of $E_{\mathbf{1}}$ are fixed with respect to the group G, the group G is cyclic and the representation is a scalar one. This (one dimensional) representation is dual to the representation of the group G on the one-dimensional space generated by any linear function. The case when there are no more components in $\mathcal{D}$, i.e. if the resolution is achieved by the first blow-up, is trivial. Otherwise let us consider a maximal component E_σ among those components E_τ of the exceptional divisor for which $G_\tau = G$ and the corresponding curvette is smooth. (The last condition can be easily detected from the resolution graph.) The smooth part $\overset{\bullet}{E}_\sigma$ of this component contains a special point x with $G_x = G$ (or all the points of $\overset{\bullet}{E}_\sigma$ are such that $G_x = G$). The point(s) from $\overset{\bullet}{E}_\sigma$ with $G_x = G$ bring(s) into the decomposition of the Poincaré series $P^G_{\{\nu_i\}}(\underline{t})$ a factor of the form $(1 - \underline{t}^{\underline{M}})^{-[G/G]_\alpha}$. The ($G$-equivariant) curvette L at the described special point of the divisor is smooth. Therefore the representation of G on the one-dimensional space generated by a G-equivariant equation of L coincides with the representation on the space generated by a linear function. Let us take all factors of the form $(1 - \underline{t}^{\underline{M}})^{-[G/G]_\alpha}$ in the decomposition of the Poincaré series $P^G_{\{v_i\}}$. For each of them, the exponent $\underline{M}$ determines the corresponding component of the exceptional divisor and therefore the topological type of the corresponding curvettes. The factor which corresponds to a component with a smooth curvette gives us the representation α on the space generated by a linear function.

Now assume that there are two special points on the first component of the resolution. Without loss of generality we can assume that they correspond to the coordinate axis $\{x = 0\}$ and $\{y = 0\}$. The representation of the group G on $\mathbb{C}^2$ is defined by its action on the linear functions x and y. For each of them this action can be recovered from a factor of the form described above just in the same way. Moreover, a factor, which determines the action of the group G on the function x, corresponds to a component of the exceptional divisor from the tail emerging from the point $\{x = 0\}$.

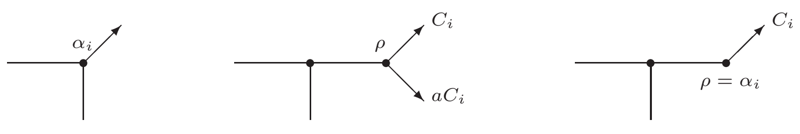
Now let G be an arbitrary (not necessarily abelian) group. For an element $g \in G$ consider the action of the cyclic group $\langle g \rangle$ generated by g on $\mathbb{C}^2$. One can see that the G-equivariant Poincaré series $P^G_{\{v_i\}}(\underline{t})$ determines the $\langle g \rangle$-

Poincaré series $P^{\langle g\rangle}_{\{v_i\}}(\underline{t})$ just like in [5, Proposition 2]. This implies that the G-equivariant Poincaré series determines the representation of the subgroup $\langle g\rangle$. (Another way is to repeat the arguments above adjusting them to the subgroup $\langle g\rangle$.) Therefore the G-Poincaré series $P^{G}_{\{v_i\}}(\underline{t})$ determines the value of the character of the G-representation on $\mathbb{C}^2$ for each element $g \in G$ and thus the representation itself. Special points of the G-action on the first component $E_{\mathbf{1}}$ of the exceptional divisor correspond to some abelian subgroups H of G. For each such subgroup H there are two special points corresponding to different one-dimensional representations of H. Again the construction above for an abelian group permits to identify tails of the dual resolution graph with these two points. □

Let $\{C_i\}$, $i = 1, \ldots, r$, be a collection of irreducible curve singularities in $(\mathbb{C}^2, 0)$ such that it does not contain curves from the same G-orbit and it does not contain a smooth curve invariant with respect to a non-trivial element of G whose action on $\mathbb{C}^2$ is not a scalar one. Let $\{\nu_i\}$ be the corresponding collection of valuations. Let $G_i \subset G$ be the isotropy group of the branch C_i, $1 \le i \le r$.

Theorem 2 *The G-equivariant Poincaré series $P^{G}_{\{\nu_i\}}(\underline{t})$ of the collection $\{\nu_i\}$ determines the weak G-equivariant topology of the collection $\{\nu_i\}$ of curve valuations.*

Proof. The minimal resolution graph Γ of the plane curve singularity $C = \bigcup_{i=1}^{r} C_i$ is essentially the same as the graph of the divisorial valuations defined by the set of irreducible components $\{E_{\alpha_i}\}$ of the exceptional divisor such that the strict transform of C_i intersects the component E_{α_i}. Instead of the mark used for the divisor E_{α_i} (like in Figures 1 and 2 for one valuation) one puts an arrow corresponding to C_i connected to the vertex α_i. Notice that there can be several arrows connected to the same vertex, i.e. $\alpha_i = \alpha_j$ for different branches C_i, C_j. In the case of one branch the graph looks like the one in Figure 2 but the vertex marked by ν coincides with τ_g and there is an arrow connected with τ_g. The number g is equal to the number of Puiseux pairs of the curve and $m_{\sigma_i} = \bar{\beta}_i$, $0 \le i \le g$, are the elements of the minimal set of generators of the semigroup of the branch. (In particular they determine the minimal resolution graph of the curve.)



(a) Γ (b) Γ^G (c) Γ enlarged

Figure 3: The graphs Γ, Γ^G and Γ enlarged.

The same rules apply for the graph Γ^G. However Γ^G corresponds to the embedded resolution of the union of all the orbits of the branches of C. So, it is possible that, in order to achieve the minimal equivariant resolution (i.e. in order to separate all the conjugates of each one of the branches C_i), one has to add some additional blow-ups starting at the point α_i. Note that in this case some of the vertices ρ (see the notations in the proof of Theorem 1 and Figures 1 and 2) do not appear in Γ. In order to preserve the scheme and the notations from the proof of the case of divisorial valuations it is better to enlarge Γ in such a way that the new one (also denoted by Γ) is the minimal one in which all the vertices ρ are present (see Figure 3). Note that $aE_{\alpha_i} = E_{a\alpha_i}$ for $a \in G$, so in this way the (new) resolution graph Γ is just the quotient of Γ^G by the obvious action of G on Γ^G.
As in the case of divisorial valuations, for each $\delta \in \Gamma^G$ let $h_\delta = 0$, $h_\delta \in \mathcal{O}_{\mathbb{C}^2,0}$, be the equation of a curvette at the component E_δ, $m_{\delta i}$ be the value $\nu_i(h_\delta)$, $M_{\delta i} = \sum_{a\in G} m_{(a\delta)i} = \sum_{a\in G}(a^*\nu_i)(h_\delta)$ and $\underline{M}_\delta = (M_{\delta 1}, \ldots, M_{\delta r}) \in \mathbb{Z}^r_{\geq 0}$. All the $\underline{M}_\sigma$, $\sigma \in \Gamma$, are different and for $\sigma, \tau \in \Gamma^G$ $\underline{M}_\sigma = \underline{M}_\tau$ if and only if $E_\tau = aE_\sigma$ for some $a \in G$. Let $G_i \subset G$ be the isotropy group of the branch C_i, $1 \leq i \leq r$.
For $i, j \in \{1, \ldots, r\}$, $m_{\alpha_i j}$ is just the intersection multiplicity between C_i and C_j and

$$M_{\alpha_i j} = \sum_{a\in G} m_{(a\alpha_i)j} = \sum_{a\in G} (a^*\nu_j)(h_{\alpha_i}) = (C_i, \bigcup_{a\in G} aC_j) = (C_j, \bigcup_{a\in G} aC_i) = M_{\alpha_j i}\,.$$

In contrast with the case of divisorial valuations the projection formula is different from the one for divisorial valuations formulated at the beginning of the proof of Theorem 1. Instead of it one has the following one: For $i_0 \in \{1, \ldots, r\}$ one has

$$P^G_{\{\nu_i\}}(\underline{t})_{|_{t_{i_0}=1}} = (1 - \underline{t}^{\underline{M}_{\alpha_{i_0}}})^{[G/G_{i_0}]_{\alpha_{i_0}}}_{|_{t_{i_0}=1}} P^G_{\{\nu_i\}_{i\neq i_0}}(t_1, \ldots, t_{i_0-1}, t_{i_0+1}, \ldots, t_r)\,. \quad (4)$$

(This can be easily deduced from (2).) Using (4) repeatedly one also has:

$$P^G_{\{\nu_i\}}(\underline{t})_{|_{t_i=1, i\neq i_0}} = \prod_{i\neq i_0} (1 - t_{i_0}^{M_{\alpha_i i_0}})^{[G/G_i]_{\alpha_i}} P^G_{\nu_{i_0}}(t_{i_0})\,. \quad (5)$$

Equations (4) and (5) imply that in order to describe inductively the minimal G-resolution graph Γ^G one has to detect the binomial $(1-\underline{t}^{\underline{M}_{\alpha_{i_0}}})$ corresponding to some i_0 from the G-equivariant Poincaré series and also the intersection multiplicities of C_{i_0} with the other branches of C. As in the divisorial case, the necessary information about the G-equivariant resolution graph can be restored from the ρ-reduction $\rho P^G_{\{\nu_i\}}(\underline{t})$ of the Poincaré series $P^G_{\{\nu_i\}}(\underline{t})$ to the ring $A(G)[[t_1, \ldots, t_r]]$. From the factorization given in (2) one can write $\rho P^G_{\{\nu_i\}}(\underline{t}) = \prod_{\sigma\in\Gamma}(1-\underline{t}^{\underline{M}_\sigma})^{s_\sigma}$, where $s_\sigma \in A(G)$. Note that the multiplicity s_σ may be equal to zero, i.e. the binomial factor corresponding to σ may be absent.

The determination of the G-equivariant resolution graph from the series $\rho P_\nu^G(t)$ for one branch almost repeats the one described for one divisorial valuation, e.g. the semigroup is the same as the one of the divisorial valuation defined by the component E_{τ_g} of the exceptional divisor. So, let us assume $r > 1$ and let us fix $j, k \in \{1, \ldots, r\}$. The separation point $s(\alpha_j, \alpha_k) \in \Gamma^G$ of α_j and α_k is defined by the condition $[\mathbf{1}, \alpha_j] \cap [\mathbf{1}, \alpha_k] = [\mathbf{1}, s(\alpha_j, \alpha_k)]$. Here $[\mathbf{1}, \sigma]$ is the geodesic in the dual graph Γ^G joining the first vertex $\mathbf{1}$ with the vertex σ. Now, let us define the separation vertex $s(\alpha_j, k)$ of C_j and GC_k as the maximun of $s(\alpha_j, a\alpha_k)$ for $a \in G$. Note that, if $a \in G$ then $s(a\alpha_j, k) = as(\alpha_j, k) \in \Gamma^G$ so $s(j, k) = s(\alpha_j, k)$ is a well defined vertex of the graph Γ. We refer to it as the separation vertex of C_i and C_j in Γ.
The ratio $M_{\sigma j}/M_{\sigma k}$ is constant for σ in $[\mathbf{1}, s(j, k)]$ and is a strictly increasing function for $\sigma \in [s(i, j), \alpha_j] \subset \Gamma$ as well as in the geodesic $[as(j, k), a\alpha_j] \subset \Gamma^G$ for $a \in G$. Notice that for $\sigma \notin \bigcup_{a \in G} ([\mathbf{1}, a\alpha_j] \cup [\mathbf{1}, a\alpha_k])$ the ratio $M_{\sigma j}/M_{\sigma k}$ is equal to $M_{\sigma' j}/M_{\sigma' k}$ where σ' is the vertex such that

$$[\mathbf{1}, \sigma'] = \max_{a \in G} \{([\mathbf{1}, a\alpha_j] \cup [\mathbf{1}, a\alpha_k]) \cap [\mathbf{1}, \sigma]\} \ .$$

Let $\sigma \in \Gamma$ be such that the exponent $\underline{M}_\sigma$ is a maximal one among the set of exponents $\underline{M}_\tau$ appearing in the factorization

$$\rho P^G_{\{\nu_i\}}(\underline{t}) = \prod_{\tau \in \Gamma \ , \ s_\tau \neq 0} (1 - \underline{t}^{\underline{M}_\tau})^{s_\tau} \ . \tag{6}$$

(Here we use the partial order $\underline{M} = (M_1, \ldots, M_r) \leq \underline{M}' = (M_1', \ldots, M_r')$ if and only if $M_i \leq M_i'$ for all $i = 1, \ldots, r$.) Note that in this case the corresponding factor has positive cardinality and there exists an index $j \in \{1, \ldots, r\}$ such that $\alpha_j = \sigma$.
Let $A \subset \{1, \ldots, r\}$ be the set of indices j such that $M_{\sigma j}/M_{\sigma k} \geq M_{\tau j}/M_{\tau k}$ for all $k \in \{1, \ldots, r\}$ and all $\tau \in \Gamma^G$ such that the binomial $(1 - \underline{t}^{\underline{M}_\tau})$ appears in (6), i.e. $s_\tau \neq 0$. From the comments above it is clear that all indices j such that $\alpha_j = \sigma$ belong to A, however A could contain some other indices ℓ such that $\alpha_\ell \neq \sigma$.
Let us assume that there exists $\ell \in A$ such that $\alpha_\ell \neq \sigma$. The behaviour of the ratios $M_{\tau\ell}/M_{\tau k}$ along $[1, \alpha_\ell]$ described above implies that $\sigma \in [1, \alpha_\ell]$. By definition of the set A, for any $\tau \in [\sigma, \alpha_\ell]$, $\tau \neq \sigma$, the binomial $(1 - \underline{t}^{\underline{M}_\tau})$ does not appear in (6), i.e. $s_\tau = 0$, in particular $\chi(\overset{\circ}{E}_\tau) = 0$. As a consequence, $\alpha_\ell < \sigma$ and α_ℓ is the end point σ_g on the dual graph of C_j (here $j \in A$ such that $\alpha_j = \sigma$). In this case one has $M_{\sigma\ell} < M_{\sigma j}$ and one can distinguish ℓ by this condition. Note that if such an $\ell \in A$ exists then it is unique.
Let $i_0 \in A$ be such that $M_{\sigma i_0} \geq M_{\sigma j}$ for all $j \in A$. Then $\alpha_{i_0} = \sigma$ and the factor $(1 - \underline{t}^{\underline{M}_{\alpha_{i_0}}})^{[G/G_{i_0}]}$ appears in the factorization (6). Thus, the projection formula permits to recover the G-equivariant resolution graph by induction.
As in Theorem 1 one has to show that the Poincaré series $P^G_{\{\nu_i\}}(\underline{t})$ determines the representation of G on $\mathbb{C}^2$, and the correspondence between “tails” of the

resolution graph. The proof in this case does not differ from the one made in Theorem 1 for divisorial valuations since the collection $\{C_i\}$ does not contains smooth curves invariant with respect to a non-trivial element of G whose action is not a scalar one. □

References

[1] Campillo A., Delgado F., Gusein-Zade S.M. The Alexander polynomial of a plane curve singularity via the ring of functions on it. Duke Math. J., v.117, no.1, 125–156 (2003).

[2] Campillo A., Delgado F., Gusein-Zade S.M. The Alexander polynomial of a plane curve singularity and integrals with respect to the Euler characteristic. Internat. J. Math., v.14, no.1, 47–54 (2003).

[3] Campillo A., Delgado F., Kiyek K. Gorenstein property and symmetry for one-dimensional local Cohen–Macaulay rings. Manuscripta Mathematica, v.83, no.3–4, 405–423 (1994).

[4] Campillo A., Delgado F., Gusein-Zade S.M. On Poincaré series of filtrations on equivariant functions of two variables. Mosc. Math. J., v.7, no.2, 243–255 (2007).

[5] Campillo A., Delgado F., Gusein-Zade S.M. Equivariant Poincaré series of filtrations. Rev. Mat. Complut., v.26, no.1, 241–251 (2013).

[6] Campillo A., Delgado F., Gusein-Zade S.M. Equivariant Poincaré series of filtrations and topology. Ark. Mat., v.52, no.1, 43–59 (2014).

[7] Campillo A., Delgado F., Gusein-Zade S.M. An equivariant Poincaré series of filtrations and monodromy zeta functions. Rev. Mat. Complut., v.28, no.2, 449–467 (2015).

[8] Gusein-Zade S.M., Delgado F., Campillo A. Universal abelian covers of rational surface singularities, and multi-index filtrations. Funktsional. Anal. i Prilozhen., v.42, no.2, 3–10 (2008); translation in Funct. Anal. Appl. v. 42, no.2, 83–88 (2008).

[9] Gusein-Zade, S.M., Luengo I., Melle-Hernández A. An equivariant version of the monodromy zeta function. Geometry, topology, and mathematical physics, 139–146, Amer. Math. Soc. Transl. Ser. 2, v.224, Amer. Math. Soc., Providence, RI, 2008.

[10] Gusein-Zade, S.M., Luengo I., Melle-Hernández A. On an equivariant version of the zeta function of a transformation. Arnold Journal of Mathematics, v.1, no.2, 127–140 (2015).

[11] Némethi, A. The Seiberg-Witten invariants of negative definite plumbed 3-manifolds. J. Eur. Math. Soc., v.13, no.4, 959–974 (2011).

A. Campillo
IMUVA (Instituto de Investigación en Matemáticas)
Universidad de Valladolid
Paseo de Belén, 7
47011 Valladolid, Spain
campillo@agt.uva.es

F. Delgado
IMUVA (Instituto de Investigación en Matemáticas)
Universidad de Valladolid
Paseo de Belén, 7
47011 Valladolid, Spain
fdelgado@agt.uva.es

S. M. Gusein-Zade
Moscow State University
Faculty of Mathematics and Mechanics
Moscow, GSP-1, 119991, Russia
sabir@mccme.ru

Free Curves on Varieties

Frank Gounelas

Received: June 26, 2014
Revised: January 30, 2016

Communicated by Thomas Peternell

Abstract. We study various generalisations of rationally connected varieties, allowing the connecting curves to be of higher genus. The main focus will be on free curves $f : C \to X$ with large unobstructed deformation space as originally defined by Kollár, but we also give definitions and basic properties of varieties X covered by a family of curves of a fixed genus g so that through any two general points of X there passes the image of a curve in the family. We prove that the existence of a free curve of genus $g \geq 1$ implies the variety is rationally connected in characteristic zero and initiate a study of the problem in positive characteristic.

2010 Mathematics Subject Classification: 14M20, 14M22, 14H10.

1. Introduction

Let k be an algebraically closed field. A smooth projective rationally connected variety, originally defined in [Cam92] and [KMM92], is a variety such that through every two general points there passes the image of a rational curve. In characteristic zero this is equivalent to the notion of a separably rationally connected variety, given by the existence of a rational curve $f : \mathbb{P}^1 \to X$ such that $f^*\mathscr{T}_X$ is ample. In characteristic p, however, one has to distinguish between these two notions. Deformations of a morphism $f : \mathbb{P}^1 \to X$ are controlled by the sheaf $f^*\mathscr{T}_X$, hence studying positivity conditions of this bundle is intimately tied to deformation theory and the existence of many rational curves on X. Rationally connected varieties have especially nice properties and an introduction to the theory is contained in [Kol96] and [Deb01]. Note in particular the important theorem of Graber-Harris-Starr [GHS03] (and de Jong-Starr [dJS03] in positive characteristic) which we will make repeated use of throughout this paper, which says that a separably rationally connected fibration over a curve admits a section. An equivalent statement in characteristic zero is that the maximal rationally connected (MRC) quotient $R(X)$ is not uniruled (see [Kol96, IV.5.6.3]), although this can fail in positive characteristic.

In this paper we study various ways in which a variety can be connected by higher genus curves. After an introductory section with auxiliary results on vector bundles on curves and Frobenius, we consider first varieties which admit a morphism from a family of curves of fixed arithmetic genus g whose product with itself dominates the product of the variety with itself and call these varieties "genus g connected", generalising the notion of there being a rational curve through two general points. We also consider C-connected varieties, where there exists a family $C \times U \to X$ of a single smooth genus g curve C such that $C \times C \times U \to X \times X$ is dominant. Mori's Bend and Break result allows us to produce rational curves going through a fixed point given a higher genus curve which has large enough deformation space. For example, in Proposition 3.6 as an easy corollary, we show that over any characteristic, if for any two general points of a smooth projective variety X with $\dim X \geq 3$ there passes the image of a morphism from a fixed curve C of genus g, then X is uniruled. This fails for surfaces, where an example is provided.

A stronger condition than the aforementioned is the existence of a morphism from a curve which deforms a lot without obstructions, as discussed for separably rationally connected varieties above. Namely, for $f : C \to X$ a morphism to a variety X where C is of any genus g, Kollár [Kol96] defines f to be free if $f^*\mathscr{T}_X$ is globally generated as a vector bundle on C and also $H^1(C, f^*\mathscr{T}_X) = 0$. In the case of genus $g = 0$ one must distinguish between free and very free curves. Geometrically, the former implies that $f : \mathbb{P}^1 \to X$ deforms so that its image covers all points in X (hence X is uniruled) whereas the latter that it can do so even fixing a point $x \in X$ (X rationally connected). If $g \geq 1$, however, after defining an r-free curve to be one which deforms keeping any r points fixed, we show that the notions of the existence of a free (0-free) and very free (1-free) curve coincide and in fact are equivalent with the existence of a curve $f : C \to X$ such that $f^*\mathscr{T}_X$ is ample.

Theorem. *(see 5.5) Let X be a smooth projective variety and C a smooth projective curve of genus $g \geq 1$ over an algebraically closed field k. Then for any $r \geq 0$, there exists an $f : C \to X$ which is r-free if and only if there exists a morphism $f' : C \to X$ such that $f'^*\mathscr{T}_X$ is ample.*

Work of Bogomolov-McQuillan (see [BM01], [KSCT07]) on foliations which restrict to an ample bundle on a smooth curve sitting inside a complex variety X shows that the leaves of such a foliation are not only algebraic but in fact have rationally connected closures. From the above, one deduces this result in the case of the foliation $\mathscr{F} = \mathscr{T}_X$, complementing the currently known connections between existence of curves with large deformation space and rationally connected varieties (cf. the uniruledness criterion of Miyaoka [Miy87]). Our proof emphasises the use of free curves and C-connected varieties, in particular with a view towards similar results in positive characteristic.

Theorem. *(see 5.2) Let X be a smooth projective variety over an algebraically closed field of characteristic zero and let $f : C \to X$ be a smooth projective curve of genus $g \geq 1$ such that $f^*\mathscr{T}_X$ is globally generated and $H^1(C, f^*\mathscr{T}_X) = 0$. Then X is rationally connected.*

In the sixth section we study the particular case of elliptically connected varieties (i.e. genus one connected varieties) where, even allowing a covering family of genus 1 curves to vary in moduli, one can prove the following theorem.

Theorem. *(Theorem 6.2) Let X be a smooth projective variety over an algebraically closed field of characteristic zero. Then the following two statements are equivalent*

(1) *There exists $\mathcal{C} \to U$ a flat projective family of irreducible genus 1 curves with a map $\mathcal{C} \to X$ such that $\mathcal{C} \times_U \mathcal{C} \to X \times X$ is dominant.*
(2) *X is either rationally connected or a rationally connected fibration over a curve of genus one.*

In positive characteristic, at this point we have not been able to prove that the existence of a higher genus free curve implies the existence of a very free rational curve (which would mean that X is separably rationally connected). We work however in this direction, establishing this result in dimensions two (with a short discussion about dimension three) and furthermore by studying algebraic implications of the existence of a free higher genus curve, such as the vanishing of pluricanonical forms and triviality of the Albanese variety. In the final section we give an example of a threefold in characteristic p whose MRC quotient is rationally connected and which has infinite fundamental group.

The study of rational curves on varieties is an important and active area of research, and shedding light on the existence of rational curves coming from the deformation theory of higher genus curves is a theme explored in a variety of sources, for example the minimal model program or [BDPP13]. Aside from the unresolved difficulties arising in positive characteristic, the author expects uniruledness and rational connected results of the type presented in this article to be of use in moduli theory.

acknowledgements. The contents of this paper are from the author's thesis under the supervision of Victor Flynn, whom I would like to thank for his continuous encouragement. I am indebted to Damiano Testa for the many hours spent helping with the material of this paper and to Johan de Jong not only for the hospitality at Columbia University but also for helping improve the contents of this paper. I would also like to thank Jason Starr and Yongqi Liang for comments, János Kollár for pointing out a similar construction to that in the last section and Mike Roth for showing me how abelian surfaces are C-connected. The anonymous referee's numerous suggestions and corrections also significantly improved this paper. This research was completed under the support of EPSRC grant number EP/F060661/1 at the University of Oxford.

2. Ample vector bundles and Frobenius

We begin with some results concerning positivity of vector bundles on curves. Recall that a locally free sheaf $\mathscr{E}$ on a scheme X is called ample if $\mathscr{O}_{\mathbb{P}(\mathscr{E})}(1)$ has this property. Equivalent definitions involving global generation of $\mathscr{F}\otimes S^n(\mathscr{E})$ for $\mathscr{F}$ a coherent sheaf and n large enough, and also cohomological vanishing criteria can be found in [Har66]. Ampleness on curves can be checked using various criteria such as the following.

Lemma 2.1. *Let C be a smooth projective curve of genus $g\geq 2$ over an algebraically closed field of characteristic zero and $\mathscr{E}$ a locally free sheaf on C such that $H^1(C,\mathscr{E})=0$. It follows that $\mathscr{E}$ is ample.*

Proof. From [Har71, Theorem 2.4], it suffices to show that every non-trivial quotient locally free sheaf of $\mathscr{E}$ has positive degree. Let $\mathscr{E}\to\mathscr{E}'\to 0$ be a quotient. From the long exact sequence in cohomology we see that $H^1(C,\mathscr{E}')$ is also 0. From the Riemann-Roch formula $\deg\mathscr{E}'=h^0(C,\mathscr{E}')+(\operatorname{rk}\mathscr{E}')(g-1)$ and since $g\geq 2$ we deduce that $\deg\mathscr{E}'>0$. $\square$

Note that Hartshorne's ampleness criterion only works in characteristic zero. More generally, over any characteristic if we further assume that our locally free sheaf is globally generated then the same result holds so long as the genus is at least one.

Proposition 2.2. *Let C be a smooth projective curve of genus $g\geq 1$ over an algebraically closed field k and $\mathscr{E}$ a globally generated locally free sheaf on C such that $H^1(C,\mathscr{E})=0$. Then $\mathscr{E}$ is ample.*

Proof. Since $\mathscr{E}$ is globally generated, there exists a positive integer n such that $\mathscr{O}_C^{\oplus n}\to\mathscr{E}\to 0$ is exact. This gives (see [Har77, ex. II.3.12]) a closed immersion of the respective projective bundles $\mathbb{P}(\mathscr{E})\hookrightarrow\mathbb{P}^{n-1}_C$. By projecting onto the first factor we have the following diagram

$$\begin{tikzcd}
\mathbb{P}(\mathscr{E}) \arrow[r, hook, "i"] \arrow[dr, "\pi"'] & \mathbb{P}^{n-1}\times C \arrow[r, "\mathrm{pr}_1"] \arrow[d, "\mathrm{pr}_2"] & \mathbb{P}^{n-1} \\
 & C &
\end{tikzcd}$$

and from [Har77, II.5.12] we have $\mathrm{pr}_1^*\mathscr{O}_{\mathbb{P}^{n-1}}(1)=\mathscr{O}_{\mathbb{P}^{n-1}_C}(1)$. Also, since i is a closed immersion it follows that $i^*\mathscr{O}_{\mathbb{P}^{n-1}_C}(1)=\mathscr{O}_{\mathbb{P}^{n-1}_C}(1)|_{\mathbb{P}(\mathscr{E})}=\mathscr{O}_{\mathbb{P}(\mathscr{E})}(1)$ which concludes that $i^*\mathrm{pr}_1^*\mathscr{O}_{\mathbb{P}^{n-1}}(1)=\mathscr{O}_{\mathbb{P}(\mathscr{E})}(1)$. To show that $\mathscr{E}$ is an ample locally free sheaf on C it is enough to show that this invertible sheaf is ample. Since we know that $\mathscr{O}_{\mathbb{P}^{n-1}}(1)$ is ample though, it is sufficient to show that $i\circ\mathrm{pr}_1$ is a finite morphism. Since it is projective, we need only show that it is quasi-finite. Hence assuming that the fibre of $i\circ\mathrm{pr}_1$ over a general point $p\in\mathbb{P}^{n-1}$ is not finite, it must be the whole of C. We now embed this fibre $j:C\to\mathbb{P}(\mathscr{E})$ as a section to π and pull back the surjection $\pi^*\mathscr{E}\to\mathscr{O}_{\mathbb{P}(\mathscr{E})}(1)$ via j, obtaining $j^*\mathscr{O}_{\mathbb{P}(\mathscr{E})}(1)$ as a quotient of $j^*\pi^*\mathscr{E}=\mathscr{E}$ (see [Har77, II.7.12]).

However $\mathrm{pr}_1 \circ i \circ j : C \to \mathbb{P}^{n-1}$ is a constant map so $j^*\mathscr{O}_{\mathbb{P}(\mathscr{E})}(1) = \mathscr{O}_C$. Taking cohomology of the corresponding short exact sequence given by this quotient, we obtain a contradiction since $H^1(C, \mathscr{E}) = 0$ whereas $H^1(C, \mathscr{O}_C)$ is not trivial for $g \geq 1$. □

In Proposition 2.4 below we will prove that given an ample bundle on a curve in positive characteristic, then after pulling back by Frobenius, we can make this bundle be globally generated and have vanishing first cohomology.

Lemma 2.3. *Let C be a smooth projective curve over an algebraically closed field k, $d \geq 0$ an integer and $\mathscr{E}$ a locally free sheaf on C. If $H^1(C, \mathscr{E}(-D)) = 0$ for all effective divisors D of fixed degree d then for $d' < d$ it follows that $H^1(C, \mathscr{E}(-D')) = 0$ and $\mathscr{E}(-D')$ is globally generated for all effective divisors D' of degree d'.*

Proof. The first result follows from the short exact sequence

$$0 \to \mathscr{E}(-D' - R) \to \mathscr{E}(-D') \to \mathscr{E}(-D')|_R \to 0$$

where R is an effective divisor of degree $d - d'$. For the second, let $p \in C$. From the first part we have $H^1(C, \mathscr{E}(-D' - p)) = 0$ since $D' + p$ is an effective divisor of degree $d' + 1 \leq d$ so the following sequence is exact

$$0 \to H^0(C, \mathscr{E}(-D' - p)) \to H^0(C, \mathscr{E}(-D')) \to \mathscr{E}(-D') \otimes k(p) \to 0.$$

Hence $\mathscr{E}(-D')$ is globally generated at p and the result follows. □

A partial converse to Proposition 2.2 in characteristic p is given in [KSCT07, Proposition 9], using $\mathbb{Q}$-twisted vector bundles as in [Laz04, II.6.4]. We prove the following different version of this result.

Proposition 2.4. *Let C be a smooth projective curve of genus g over an algebraically closed field k of characteristic p and let $\mathscr{E}$ be an ample locally free sheaf on C. Let $B \subset C$ be a closed subscheme of length b and ideal sheaf $\mathscr{I}_B$. Then there exists a positive integer n such that $H^1(C^{(n)}, F_n^*\mathscr{E} \otimes \mathscr{I}_B) = 0$ and $F_n^*\mathscr{E} \otimes \mathscr{I}_B$ is globally generated on $C^{(n)}$ where $F_n : C^{(n)} \to C$ the n-fold composition of the k-linear Frobenius morphism.*

Proof. We proceed by induction. First, assume we can write $\mathscr{E}$ as an extension

$$0 \to \mathscr{M} \to \mathscr{E} \to \mathscr{Q} \to 0$$

where $\mathscr{M}$ is an ample line bundle. If $\mathscr{Q}$ is not torsion free, consider the saturation of $\mathscr{M}$ in $\mathscr{E}$ instead and take $\mathscr{Q}$ as that quotient. Since $\mathscr{E}$ is ample, so is its quotient $\mathscr{Q}$. Note also that the rank of $\mathscr{Q}$ is one less than that of $\mathscr{E}$ and that if we can prove the result for $\mathscr{Q}$ then we will have it for $\mathscr{E}$ too by considering cohomology of the appropriate exact sequences. We thus reduce to the case of $\mathscr{E} = \mathscr{L}$ an invertible sheaf of positive degree (since it is ample). An invertible sheaf $\mathscr{L}$ pulls back under the n-fold composition of the linear Frobenius morphism to an invertible sheaf $F_n^*\mathscr{L}$ of degree $p^n \deg \mathscr{L}$. To show that $H^1(C^{(n)}, F_n^*\mathscr{L} \otimes \mathscr{I}_B) = 0$, it is equivalent by Serre duality to show that $\mathrm{Hom}_{C^{(n)}}(F_n^*\mathscr{L}, \mathscr{O}_{C^{(n)}}(B)\otimes\omega_{C^{(n)}}) = 0$. Since the invertible sheaf $\mathscr{O}_{C^{(n)}}(B)\otimes\omega_{C^{(n)}}$ has degree $b+2g-2$ and by picking n large enough, we can ensure $p^n \deg \mathscr{L} > b+2g-2$ from which we obtain $H^1(C^{(n)}, F_n^*\mathscr{L}\otimes\mathscr{I}_B) = 0$ and hence $H^1(C^{(n)}, F_n^*\mathscr{E}\otimes\mathscr{I}_B) = 0$ for a locally free sheaf of any rank.

To show that $F_n^*\mathscr{E}\otimes\mathscr{I}_B$ is globally generated, pick a point $q \in C$. Then $\mathscr{I}_B\otimes\mathscr{I}_q$ has length $b+1$ and from the discussion above $H^1(C^{(n)}, F_n^*\mathscr{E}\otimes\mathscr{I}_B\otimes\mathscr{I}_q)$ vanishes when $p^n \deg L > b+1+2g-2$ so we can just pick n large enough to fit this condition. Now, by taking the long exact sequence in cohomology of

$$0 \to F_n^*\mathscr{E}\otimes\mathscr{I}_B\otimes\mathscr{I}_q \to F_n^*\mathscr{E}\otimes\mathscr{I}_B \to (F_n^*\mathscr{E}\otimes\mathscr{I}_B)\otimes k(q) \to 0$$

we conclude that $F_n^*\mathscr{E}\otimes\mathscr{I}_B$ is globally generated.

That $\mathscr{E}$ can not be written as an extension of $\mathscr{M}$ an ample line bundle and a quotient locally free sheaf $\mathscr{Q}$ is equivalent to $H^0(C, \mathscr{E}\otimes\mathscr{M}^{-1}) = 0$. However there exists a positive integer m and an ample line bundle $\mathscr{M}_{C^{(m)}}$ on $C^{(m)}$ for which $H^0(C^{(m)}, (F_m^*\mathscr{E})\otimes\mathscr{M}_{C^{(m)}}^{-1}) \neq 0$ and we proceed as before with the sheaf $(F_m^*\mathscr{E})$. □

3. Definition of curve connectedness: Covering families

We now define various ways in which a variety can be covered by curves, generalising the notion of a rationally connected varieties (see [Kol96, IV]).

Definition 3.1. We say that a variety X over a field k is *connected by genus $g \geq 0$ curves* (resp. *chain connected by genus g curves*) if there exists a proper flat morphism $\mathcal{C} \to Y$, for a variety Y, whose geometric fibres are irreducible genus g curves (resp. connected genus g curves) such that there is a morphism $u : \mathcal{C} \to X$ making the induced morphism $u^{(2)} : \mathcal{C}\times_Y\mathcal{C} \to X\times_k X$ dominant.

We say X is separably (chain) connected by genus g curves if $u^{(2)}$ is smooth at the generic point. Note that the notion of separability is redundant in characteristic zero due to generic smoothness. A genus zero connected variety is rationally connected. A variety which is connected by genus one curves will be called (with a slight abuse of notation) elliptically connected. The relevant moduli spaces which we will be considering are the following. Let $\pi : \mathcal{C} \to S$ be a flat projective curve over an irreducible scheme S and let $B \subset \mathcal{C}$ be a closed subscheme that is flat and finite over S. Let $p : X \to S$ be a smooth quasi-projective scheme and $g : B \to X$ an S-morphism. The space (see [Kol96, II.1.5] and [Mor79]) $\mathrm{Hom}_S(\mathcal{C}, X, g)$ parametrises S-morphisms from $\mathcal{C}$ to X keeping the points given by g fixed. Restricting to the case where S is the spectrum of an algebraically closed field k we fix some notation of the following evaluation morphisms to be used in later sections

$$\begin{aligned} F : C\times\mathrm{Hom}(C, X, g) &\to X \\ \phi(p, f) : H^0(C, f^*\mathscr{T}_X\otimes\mathscr{I}_B) &\to f^*\mathscr{T}_X\otimes k(p) \end{aligned}$$

and similarly the double evaluation morphisms $F^{(2)}$ and $\phi^{(2)}(p,q,f)$ as in [Kol96, II.3.3]. Secondly we consider the relative moduli space of genus g degree d stable curves with base point $t : P \to X$, denoted by $\overline{\mathcal{M}}_g(X/S,d,t)$ as in [AK03] (originally [FP97]). By Bertini, we can always find a genus g such that a projective X is genus g connected, the minimal such g however is an interesting invariant of the variety. Finding higher genus covering families is an easy operation.

Lemma 3.2. *Let X be a genus g (chain) connected smooth projective variety over an algebraically closed field k. Then if $g' \geq 2g-1$, X is also genus g' (chain) connected.*

Proof. Let $\mathcal{C}/Y \to X$ be a family making X a genus g (chain) connected variety. From [AK03, Theorem 50] we have a projective algebraic space $Y' = \overline{\mathcal{M}}'_g(\mathcal{C}/Y,d)$ of finite type over Y parametrising stable families of degree d curves of genus g' over $\mathcal{C} \to Y$. The condition $g' \geq 2g-1$ coming from the Riemann-Hurwitz formula ensures that this moduli space is non-empty. From [ACG11, 12.9.2] there exists a normal scheme Z finite and surjective over Y' and a flat and proper family $\mathcal{X} \to Z$ of stable genus g curves of degree d. Restricting to a suitable open subset $W \subset Z$ parametrising irreducible curves we compose the family $\mathcal{X}|_W \to W$ with the evaluation morphism to X and the result follows. $\square$

An example of an elliptically connected variety over a non-algebraically closed field is given after the proof of Theorem 6.2. A much stronger condition is the existence of a family of curves which is constant in moduli.

Definition 3.3. We say that a variety X over a field k is *C-connected* for a curve C if there exists a variety Y and a map $u : C \times Y \to X$ such that the induced map $u^{(2)} : C \times C \times Y \to X \times X$ is dominant. If $u^{(2)}$ is also smooth at the generic point, then we say that X is *separably C-connected*.

Projective space is C-connected for every smooth projective curve C whereas an example of a C-connected variety which is not rationally connected is $C \times \mathbb{P}^n$ where $g(C) \geq 1$. To see this let $(c_1,x_1),(c_2,x_2)$ be any two points in $C \times \mathbb{P}^n$ and let $f : C \to \mathbb{P}^n$ a morphism which sends $c_i \mapsto x_i$. Considering the graph of f in $C \times \mathbb{P}^n$ we have found a curve isomorphic to C which goes through our two points. Using parts (3) and (4) from Lemma 3.4 below, the result follows. More generally, examples can also be constructed from Proposition 3.5 below. The following are mostly straight forward generalisations of various results in [Kol96, IV.3].

Lemma 3.4. *The following statements hold for a variety X over a field k and C a smooth projective curve.*

(1) *If X is genus g connected and $X \dashrightarrow Y$ a dominant rational map to a proper variety Y, then Y is also genus g connected. The same holds if X is C-connected.*

(2) *A variety X is C-connected if and only if there is a variety W, closed in* $\mathrm{Hom}(C,X)$ *such that* $u^{(2)} : C \times C \times W \to X \times X$ *is dominant.*
(3) *If X is defined over a field k and K/k is an extension of fields, then* $X_K := X \times_k K$ *is C-connected if and only if X_k is.*
(4) *A variety X over an uncountable algebraically closed field is C-connected if and only if for all very general $x_1, x_2 \in X$ there exists a morphism $C \to X$ which passes through x_1, x_2.*
(5) *A variety X over an uncountable algebraically closed field is genus g connected if and only if for all very general $x_1, x_2 \in X$ there exists a smooth irreducible genus g curve containing them.*
(6) *Being rationally or elliptically connected is closed under connected finite étale covers of varieties.*

PROOF. To prove (1), let $u : \mathcal{C}/M \to X$ be a family making X genus g connected and denote by $u' : \mathcal{C}/M \dashrightarrow Y$ the composition. Restricting u' to the generic fibre $\mathcal{C}_{k(M)}$ we have a rational map $\phi : \mathcal{C}_{k(M)} \dashrightarrow Y$. Since Y is proper, by the valuative criterion of properness we can extend ϕ to a morphism $\phi : \mathcal{C}_{k(M)} \to Y$. By spreading out to an open subset $M' \subseteq M$ (see [DG67, IV_3 8.10.5] for properness and 11.2.6 for flatness of the family) we obtain a family $\mathcal{C}|_{M'} \to M'$ which makes Y also genus g connected.

Since being C-connected or connected by genus g curves is a birational property, we may assume by compactifying that X is projective. For (2), consider $\mathrm{Hom}(C,X) = \cup R_i$ the decomposition into irreducible components. One direction of the statement is obvious, whereas for the other let $C \times W \to W$ be a family which makes X a C-connected variety. If $u_i : C \times R_i \to X$ is the evaluation morphism, then for some i there is a morphism $h : W \to R_i$ such that $h(w) = [C_w \to X]$ for general $w \in W$. This implies that $u_i^{(2)} : C \times C \times R_i \to X \times X$ is also dominant. For one direction of (3), pullback by $\operatorname{Spec} K \to \operatorname{Spec} k$. For the other, if X_K is C_K-connected then from (2) there is a positive integer d such that the evaluation morphism $\mathrm{ev}_K^d : C_K \times C_K \times \mathrm{Hom}_d(C_K, X_K) \to X_K \times X_K$ is dominant. Because of the universal property of the Hom-scheme, we have that $\mathrm{Hom}(C,X) \times_k K = \mathrm{Hom}(C_K, X_K)$ and $(\mathrm{ev}^d)_K = \mathrm{ev}_K^d$ so ev^d is also dominant.

If through every two very general points there passes the image of C under some morphism, then the map $u^{(2)} : C \times C \times \mathrm{Hom}(C,X) \to X \times X$ is dominant. Since $\mathrm{Hom}(C,X)$ has at most countably many irreducible components the restriction of $u^{(2)}$ to at least one of the components R_i must be dominant, which proves (4). Similarly for (5) working instead with the Kontsevich moduli of curves $\mathcal{M}_{g,1}(X) \to \mathcal{M}_{g,0}(X)$ the result follows. For (6), the proof for rationally connected varieties is contained in [Deb01, 4.4.(5)]. Let $\mathcal{C} \to U$ be a family which makes X elliptically connected and let $X' \to X$ be a connected finite étale cover. Consider the pullback diagram and $\mathcal{C}' \to U' \to U$ the Stein factorisation

$$\begin{tikzcd}
& \mathcal{C}' = \mathcal{C} \times_X X' \arrow[r] \arrow[dl] \arrow[d] & X' \arrow[d] \\
U' \arrow[dr] & \mathcal{C} \arrow[r] \arrow[d] & X \\
& U. &
\end{tikzcd}$$

After possibly restricting U' to the open subset of curves in $\mathcal{C}'$ which are irreducible, the family $\mathcal{C}' \to U'$ makes X' elliptically connected. $\square$

Proposition 3.5. *Let X be a smooth projective variety over an algebraically closed field k and $f : X \to C$ a flat morphism to a smooth projective curve whose geometric generic fibre is separably rationally connected. Then X is C-connected.*

Proof. From [dJS03], there is a section $\sigma : C \to X$ to f. Now from [KMM92, Theorem 2.13] we can find a section to f passing through any two points in different smooth fibres over C, hence we can find a copy of C passing through two general points. The result now follows from Lemma 3.4 parts (4) and (5) above after possibly passing to an uncountable extension K/k. $\square$

We now come to the main theme of this paper, which is that varieties covered by higher genus curves in a strong sense are also covered by rational curves. This is illustrated in the following proposition, and continues in the next sections.

Proposition 3.6. *Let X be a C-connected variety of dimension at least 3 over an algebraically closed field k. Then X is uniruled.*

Proof. We may assume X is projective. Let $u : C \times Y \to X$ be a family such that $u^{(2)} : C \times C \times Y \to X \times X$ is dominant. We have $\dim Y + 2 \geq 2 \dim X$ and so if $\dim X \geq 3$ we obtain $\dim Y \geq 4$. Now, pick general points $x \in X, c \in C$ and denote by $Z \subset Y$ the locus of curves $u_z : C_z \to X$ such that $x = u_z(c)$ for all $z \in Z$. We have that $\dim Z \geq \dim Y - (\dim X - 1) - 1$ and so for $\dim X \geq 3$, $\dim Z \geq 1$. Since any two general points in X can be connected by the image of a C_y, it follows that Z does not get contracted to a point when mapped to $\mathrm{Hom}(C, X; c \mapsto x)$. From Bend and Break (see [Deb01, Prop. 3.1]) we obtain a rational curve through x and hence through every general point. After possibly an extension to an uncountable algebraically closed field this implies that X is uniruled (see [Deb01, Remark 4.2(5)]). $\square$

If C has genus one, the above result is also proved in Section 6, even allowing the curve C to vary in moduli and with the dimension of X assumed greater or equal to two. On the other hand, a C-connected surface does not have to be uniruled when C has genus at least two. Consider $C \subset A$ a curve in an abelian surface such that C contains the identity 0 of A and the genus of C is at least two. Consider the map $\phi : C \times C \to X$ sending (p, q) to $p - q$. If the image is one dimensional, it has to be isomorphic to C since it has to be irreducible and contains the image of $C \times \{0\}$. On the other hand, the image will be closed under the group operation, hence would have to be abelian itself, which is a contradiction. Hence ϕ is surjective, and we obtain that for any $x \in A$, there is a $(p, q) \mapsto x$, hence a morphism $C \cong C \times \{q\} \to X$ passing through x and 0 (for (q, q)). Take any two points $x, y \in A$, and consider the image of a morphism from C through 0 and the point $x - y$ that we just constructed. Translate this curve by y and obtain an image of C through x, y.

Denoting by $X \dashrightarrow R(X)$ the maximal rationally chain connected (MRC) fibration, we let $R^0(X) = X$, $R^i(X) = R(R^{i-1}X)$ and obtain a tower of MRC fibrations

$$X \dashrightarrow R^1(X) \dashrightarrow \cdots \dashrightarrow R^n(X).$$

This tower eventually stabilises, and if $R^i(X)$ is uniruled then $\dim R^{i+1}(X) < \dim R^i(X)$. In characteristic zero, we in fact have $R(X) = \ldots = R^n(X)$ (see discussion below). In positive characteristic it can be that the tower has length greater than one - see the example given in the last section of this paper.

Proposition 3.7. *Let X be a normal and proper C-connected variety over an algebraically closed field where C is a smooth projective curve. Then the tower $X \dashrightarrow R^1(X) \dashrightarrow \cdots \dashrightarrow R^n(X)$ of MRC quotients terminates in either a point, a curve or a surface.*

Proof. Let $C \times Y \to X$ be a family which makes X a C-connected variety. From Lemma 3.4 part (1) it follows that $R^i(X)$ are also C-connected. From Proposition 3.6 we obtain that $R^i(X)$ is uniruled if $\dim R^i(X) \geq 3$. This implies that $R^{i+1}(X)$ must have dimension strictly less than $R^i(X)$ and so the result follows. □

Note that if k is algebraically closed of characteristic zero then we know from [GHS03] that the MRC quotient $R(X)$ is not uniruled, so if X is C-connected of dimension at least three, $R(X)$ must be a surface, curve or point, in which case X is respectively a rationally connected fibration over a surface or curve, or a point (and so X is rationally connected). From Proposition 3.5 the converse holds too for a fibration over a curve.

Remark 3.8. As observed in [Occ06, Remark 4], if the MRC quotient of a smooth complex projective variety X is a curve, then the MRC fibration extends to the whole variety and coincides with the Albanese map.

4. Definition of curve connectedness: Free morphisms

In this section we define ways in which a morphism from a curve C to a variety X can deform enough to give a large family of morphisms from C so as to cover X. A notion studied extensively by Hartshorne [Har70] is that of a (local complete intersection) subvariety Y in a smooth projective variety X such that the normal bundle $\mathscr{N}_{Y/X}$ is ample. Hartshorne proved in [Har70, III.4] that for some $g \geq 0$ there exists a curve $C \subset X$ of genus g such that $\mathscr{N}_{C/X}$ is ample. Alternatively, Ottem [Ott12] defines an ample closed subscheme $Y \subset X$ of codimension r to be one where the exceptional divisor $\mathscr{O}(E)$ of the blowup $\mathrm{Bl}_Y X$ of X along Y is an $(r-1)$-ample line bundle in the sense that for every coherent sheaf $\mathscr{F}$ there is an integer $m_0 > 0$ such that $H^i(X, \mathscr{F} \otimes \mathscr{O}(E)^m) = 0$ for all $m > m_0$ and $i > r-1$. One can then prove that if Y is a local complete intersection subscheme of X which is ample, then the normal bundle $\mathscr{N}_{Y/X}$ is an ample bundle. We impose the following stronger positivity condition.

Definition 4.1. ([Kol96, II.3.1]) Let C be a smooth proper curve and X a smooth variety over a field k. Let $f : C \to X$ a morphism and $B \subset C$ a closed subscheme with ideal sheaf $\mathscr{I}_B$ and $g = f|_B$. The morphism f is called *free over g* if it is non-constant and one of the following two equivalent conditions is satisfied:

(1) for every $p \in C$ we have $H^1(C, f^*\mathscr{T}_X \otimes \mathscr{I}_B(-p)) = 0$ or,
(2) $H^1(C, f^*\mathscr{T}_X \otimes \mathscr{I}_B) = 0$ and $f^*\mathscr{T}_X \otimes \mathscr{I}_B$ is generated by global sections.

Note that there is also a relative version of the above definition discussed in [KSCT07].

Definition 4.2. We say that a curve $f : C \to X$ is *r-free* if for all effective divisors D of degree $r \geq 0$, $H^1(C, f^*\mathscr{T}_X \otimes \mathscr{O}_C(-D)) = 0$ and $f^*\mathscr{T}_X \otimes \mathscr{O}_C(-D)$ is generated by global sections. A 0-free curve is called *free* whereas a 1-free curve is called *very free*.

The condition of r-freeness makes formal the notion that the curve C deforms in X while keeping any general r points fixed. The following follows immediately from Lemma 2.3.

Lemma 4.3. *If $f : C \to X$ is an r-free curve then f is r'-free for all $r' \leq r$.*

In the case of $C = \mathbb{P}^1$, $f^*\mathscr{T}_X = \oplus_{i=1}^n \mathscr{O}_{\mathbb{P}^1}(a_i)$ with $a_1 \leq \ldots \leq a_n$ so it follows that $f : \mathbb{P}^1 \to X$ is r-free if and only if $a_1 \geq r$.

Remark 4.4. We should remark at this point that there do not exist complete intersection curves of large enough degree which are free on a general smooth hypersurface. For example, let X be a degree d smooth hypersurface in $\mathbb{P}^n$. Assume $d \leq n$ since otherwise X will be of general type or Calabi-Yau and will not have any free curves. Let Y_i be $n-2$ suitably general hypersurfaces in $\mathbb{P}^n$ all of degree e and let $C = X \cap_{i=1}^{n-2} Y_i$ be the resulting curve. The degree of C is de^{n-2} and the normal bundle is

$$\mathscr{N}_{C/X} = \oplus_{i=1}^{n-2} \mathscr{O}_{\mathbb{P}^n}(Y_i)|_C = \oplus_{i=1}^{n-2} \mathscr{O}_{\mathbb{P}^n}(e)|_C.$$

By adjunction, we compute

$$\deg \mathscr{T}_C = -\deg \omega_C = -d(-n-1+d+\sum_{i=1}^{n-2} e).$$

Even setting $e = 1$ to make $\deg \mathscr{T}_C$ as large as possible, and taking into account that $\deg \mathscr{N}_{C/X} = e(n-2)$, we see that $\deg \mathscr{T}_X|_C = \deg \mathscr{T}_C + \deg \mathscr{N}_{C/X}$ is not going to be positive for large values of d and n. Positivity of the degree of $\mathscr{T}_X|_C$ would be necessary for any ampleness conditions. See [Gou14] for a discussion on separable rational connectedness of Fano complete intersections.

A result of Kollár ([Kol96, II.1.8]) implies that if the dimension of X is at least 3, a general deformation of a 2-free morphism is an embedding into X. We will see (Theorem 5.5) that if the genus of C is at least one, this holds for any free morphism too. From [Kol96, II.3.2], if a family of curves mapping to a variety has a member which is free over g, then the locus of all such curves in this family is open.

LEMMA 4.5. *Let X be a smooth variety over an algebraically closed field k, $D \subset X$ a divisor and $f : C \to X$ a free morphism. If $p \in C$ then there exists a deformation $f' : C \to X$ with $f'(p) \notin D$.*

PROOF. By semicontinuity let $U \subset \operatorname{Hom}(C, X)$ be a connected open neighbourhood of $[f]$ such that $H^1(C, f_t^* \mathscr{T}_X) = 0$ for all $[f_t] \in U$. From [Mor79] it follows that the dimension of U is $h^0(C, f^* \mathscr{T}_X)$. Denote by $\mathscr{I}_p$ the ideal sheaf on C of the closed subscheme with unique point p. Since f is free, we have $H^1(C, f_t^* \mathscr{T}_X \otimes \mathscr{I}_p) = 0$ for all $[f_t] \in U$ and so by fixing a point $x \in X$ such that $p \mapsto x$, we have

$$\begin{aligned} \dim(\operatorname{Hom}(C, X; p \mapsto x) \cap U) &= h^0(C, f^* \mathscr{T}_X \otimes \mathscr{I}_p) \\ &= h^0(C, f^* \mathscr{T}_X) - \dim X \\ &= \dim U - \dim X. \end{aligned}$$

Next, denote by

$$V = \{[f_t] \in U \mid f_t(p) \in D\} = \bigcup_{x \in D} \{[f_t] \in U \mid f_t(p) = x\}$$

the subspace of all morphisms in U which send p to a point in the divisor D. It follows that

$$\begin{aligned} \operatorname{codim}(V, U) &\geq \dim U - \dim V \\ &= h^0(C, f^* \mathscr{T}_X) - (h^0(C, f^* \mathscr{T}_X) - \dim X + \dim X - 1) = 1 \end{aligned}$$

and hence there exists an $[f'] \in U \setminus V$ such that $f'(p) \notin D$. □

PROPOSITION 4.6. *Let X be a smooth variety over an algebraically closed field k and $f : C \to X$ a smooth projective curve which is free over $B \subset C$ a closed subscheme with ideal sheaf $\mathscr{I}_B$. Let $g : X \dashrightarrow Y$ be a generically smooth dominant rational map to a smooth proper variety Y. Then it follows that $f' := g \circ f : C \dashrightarrow Y$ can be deformed to a morphism free over B.*

PROOF. Deform $f : C \to X$ so that it misses the codimension 2 exceptional locus of g (from [Kol96, II.3.7]) so we can assume that the composition $g \circ f : C \dashrightarrow Y$ is in fact a non-constant morphism. Starting with the standard exact sequence of tangent bundles on X and applying f^* and tensoring with $\mathscr{I}_B$ we obtain

$$0 \to f^*\mathscr{T}_{X/Y} \otimes \mathscr{I}_B \to f^*\mathscr{T}_X \otimes \mathscr{I}_B \to (g \circ f)^*\mathscr{T}_Y \otimes \mathscr{I}_B. \tag{4.1}$$

From [Liu02, Ex. 6.2.10] this is exact on the right and we conclude. □

In the case of higher genus curves there exist genus g connected varieties which do not have a free or very free curve for all $g \geq 1$, for example consider $E \times \mathbb{P}^1$ where E is an elliptic curve. As pointed out after Definition 3.3, $E \times \mathbb{P}^1$ is E-connected yet it is not possible that there exists a morphism $f : C \to E \times \mathbb{P}^1$ from a curve C such that $f^*\mathscr{T}_{E\times\mathbb{P}^1}$ is ample since this bundle is isomorphic to $\mathscr{O}_C \oplus \mathscr{O}_C(2)$ which has a non-ample quotient $\mathscr{O}_C$. One can however prove the following proposition.

PROPOSITION 4.7. *Let X be a smooth variety over an algebraically closed field and $f : C \to X$ a very free morphism for some smooth projective curve C. Then X is separably C-connected.*

PROOF. Let $[f] \in Y \subset \mathrm{Hom}(C, X)$ be an open and smooth neighbourhood with cycle map $u : C \times Y \to X$. We first show that the evaluation map

$$\phi^{(2)}(p, q, f) : H^0(C, f^*\mathscr{T}_X) \to f^*\mathscr{T}_X \otimes k(p) \oplus f^*\mathscr{T}_X \otimes k(q)$$

is surjective for $p \neq q$ general points in C. Consider the following exact sequences of sheaves

$$0 \to f^*\mathscr{T}_X(-p-q) \to f^*\mathscr{T}_X \to (f^*\mathscr{T}_X \otimes k(p)) \oplus (f^*\mathscr{T}_X \otimes k(q)) \to 0$$
$$0 \to f^*\mathscr{T}_X(-p-q) \to f^*\mathscr{T}_X(-p) \to f^*\mathscr{T}_X(-p) \otimes k(q) \to 0$$

and note that by taking the long exact sequence in cohomology of the first, to show that $\phi^{(2)}(p, q, f)$ is surjective, we need to show that $H^1(C, f^*\mathscr{T}_X(-p-q)) = 0$. Since f is very free we have from the second sequence that $H^0(C, f^*\mathscr{T}_X(-p)) \to f^*\mathscr{T}_X(-p) \otimes k(q)$ is surjective and also that $H^1(C, f^*\mathscr{T}_X(-p)) = 0$ from which it follows that $H^1(C, f^*\mathscr{T}_X(-p-q)) = 0$. Since $\phi^{(2)}(p, q, f)$ is surjective, it follows from [Kol96, II.3.5] that $u^{(2)} : C \times C \times Y \to X \times X$ is smooth at $(p, q, [f])$. We conclude that X is separably C-connected and thus also separably connected by genus g curves. □

REMARK 4.8. It follows that in the setting above that a very free curve (or in fact even a C such that X is C-connected) has the property that it intersects non-trivially all but a finite number of divisors. This follows from the fact that we can cover an open subset by images of C, whose complement will be a proper closed subset of X and so contains a finite number of divisors.

5. PROVING UNIRULEDNESS AND RATIONAL CONNECTEDNESS

In this section we prove that the existence of a free curve of genus $g \geq 1$ is equivalent to the existence of an r-free curve of genus g for all $r \geq 1$, and that in characteristic zero this is also equivalent to the existence of a very free rational curve. This is in stark contrast to rational curves, where uniruled varieties (possessing free rational curves) are not always rationally connected (possessing very free rational curves). We begin by noting that there is another type of positive curve one can consider for a smooth projective variety X, namely $f : C \to X$ such that $f^*\mathscr{T}_X$ is ample. Note that such a curve automatically has $\mathscr{N}_{C/X}$ ample. Such curves have traditionally been studied in terms of foliations (cf. Theorem 5.3). We will also prove that the existence of a curve such that $f^*\mathscr{T}_X$ is ample is in fact equivalent to the existence of a free curve of the same genus.

PROPOSITION 5.1. *Let X be a smooth projective variety over an algebraically closed field k and $f : C \to X$ a morphism from a smooth projective curve of genus g such that $f^*\mathscr{T}_X$ ample. Then X is uniruled.*

PROOF. The proof follows the usual Mori argument so we present only a sketch (cf. Theorem 5.3). Note that if X is a curve, then since a bundle is ample if and only if its pullback under a finite morphism is ample, we obtain that $X = \mathbb{P}^1$. In characteristic zero, after spreading out over a finitely generated extension $\operatorname{Spec} S$ of $\operatorname{Spec} \mathbb{Z}$, one can reduce to any closed prime and consider the equivalent set-up in positive characteristic. After pulling back by Frobenius, Lemma 2.4 implies that there is a morphism $f_p^{(n)} : C_p \to X_p$ such that $(f_p^{(n)})^*\mathscr{T}_{X_p}$ is very free (or r-free even), where $f_p : C_p \to X_p$ the reduction of $f : C \to X$. Bend and Break now produces a rational curve passing through a general point, of bounded degree independent of p (see [Deb01, Prop. 3.5]). These are points in fibres over $\operatorname{Spec} S$ of a finite type relative moduli $\operatorname{Hom}^d_S(\mathbb{P}^1_S, \mathcal{X}/S, s)$, for $s : \operatorname{Spec} S \to \mathcal{X}$ a section specifying the general point the rational curve goes through. Hence by Chevalley's Theorem the generic fibre over $\operatorname{Spec} S$ is also non-empty, and there is a rational curve through a general point of X. $\square$

THEOREM 5.2. *Let X be a smooth projective variety over an algebraically closed field k and $f : C \to X$ a morphism from a smooth projective curve of genus g such that $f^*\mathscr{T}_X$ is ample.*

(1) *If the characteristic p of k is zero, then X is rationally connected.*
(2) *If $p > 0$ then the tower of MRC fibrations terminates with a point.*

PROOF. From 5.1, we conclude that X is uniruled, regardless of the characteristic. Denote by $\pi : X \to R(X)$ the MRC fibration ($R(X)$ is defined up to birational transformation so we may assume π is a morphism). In characteristic zero, the composition $g : C \to X \to R(X)$ again has $g^*\mathscr{T}_{R(X)}$ ample, since from the proof of 4.6 the quotient of an ample bundle is ample. So by the Graber-Harris-Starr Theorem, since $R(X)$ is uniruled by Proposition 5.1, it must be a point. In positive characteristic, it may not be the case that the composition $g : X \to R(X)$ is generically smooth, in which case $g^*\mathscr{T}_{R(X)}$ might not be ample. From Lemma 2.4 however there is a morphism $h : C' \to X$ such that $h^*\mathscr{T}_X$ is very free (here C' is a Frobenius pullback of C so has the same genus). From 4.7, X is separably C'-connected, and so by 3.7 we do obtain that the tower of MRC quotients $X \to R(X) \to \cdots R^n(X)$ ends in a point, curve or surface. If $\pi : X \to T$ where $T := R^n(X)$ is a smooth projective curve, then by Lemma 4.5, for a point $p \in C'$, we can deform h so that the image of p misses the inverse image under π of $\pi(h(p))$. Hence $\mathrm{Hom}(C', T)$ is at least one dimensional and from de Franchis' Theorem [ACG11, 8.27] it follows that T has genus zero or one. One excludes the case where C', T both of genus one, by using the fact that there are only countably many isogenies between two elliptic curves. Also, T cannot be rational since we have assumed the tower is maximal. If now $R^n(X) = S$ is a smooth projective surface, we may assume by pulling back by Frobenius from 2.4 and deforming, that there is an at least one dimensional family of morphisms sending a fixed point on C to a fixed point on S. Hence by Bend and Break [Deb01, Prop. 3.1] the surface would have to be uniruled and we are reduced to the case of a point again. ☐

Assuming ampleness and regularity of a foliation on a smooth curve in characteristic zero, results of this type have been demonstrated in the work of various people, starting with Miyaoka's uniruledness criterion [Miy87, Theorem 8.5]. A short summary of recent results follows.

Theorem 5.3. *([BM01, Theorem 0.1], [KSCT07, Theorem 1]) Let X be a normal complex projective variety and $C \subset X$ a complete curve in the smooth locus of X. Assume that $\mathscr{F} \subset \mathscr{T}_X$ is a foliation regular along C and such that $\mathscr{F}|_C$ is ample. If $x \in C$ is any point, the leaf through x is algebraic and if $x \in C$ is general then the closure of the leaf is also rationally connected.*

Using [BDPP13, Corollary 0.3], Peternell proved a weaker version of Mumford's conjecture on numerical characterisation of rationally connected varieties from which one can deduce the following theorem.

Theorem 5.4. *([Pet06, 5.4, 5.5]) Let $X/\mathbb{C}$ be a projective manifold and $C \subset X$ a possibly singular curve. If $\mathscr{T}_X|_C$ is ample then X is rationally connected. If $\mathscr{T}_X|_C$ is nef and $-K_X.C > 0$ then X is uniruled.*

The precise relation between r-free morphisms and morphisms $f : C \to X$ such that $f^*\mathscr{T}_X$ is ample is given in the following.

Theorem 5.5. *Let X be a smooth projective variety over an algebraically closed field k and $r \geq 0$ any integer. Then there exists a morphism $f : C \to X$ from a smooth projective genus $g \geq 1$ curve C such that $f^*\mathscr{T}_X$ is ample if and only if there is an r-free morphism $h : C' \to X$ from a genus g smooth projective curve C'.*

Proof. Assuming the existence of h, we obtain from Lemma 2.3 that h is also free, and so by Proposition 2.2, $h^*\mathscr{T}_X$ is ample. If $f^*\mathscr{T}_X$ is ample, one needs to separate between characteristic $p > 0$ or equal to zero. In the former case, as in the proof of 5.1 we get $h : C' \to X$ (here again C' is a Frobenius pullback of C so of genus g) which is r-free. When the characteristic is zero, X will be rationally connected from 5.2. The idea now is to attach many very free rational curves to C, apply standard smoothing of combs techniques and prove that the resulting general smooth deformations of the comb will be r-free genus g curves (cf [Kol96, II.7.10]). This proceeds as follows. Assemble a comb $D = C \cup \cup_{i=1}^m C_i$ with m rational teeth that are $(r+1)$-free like in [Kol96, II.7]. For m large enough, D is smoothable to a flat proper family $Y \to T$ where the general fibre is isomorphic to C, the central fibre is a subcomb of D with a large number of teeth depending on $C \subset X$ and m, and there is a morphism $F : Y \to X$ which extends $D \to X$. To show that the general nearby fibre $f_t : Y_t \to X$ is r-free, it suffices to show that $H^1(Y_t, f_t^* \mathscr{T}_X(-\sum_{i=0}^r p_i))$ for $p_0, p_1, \ldots, p_r$ any points on $Y_t \subset X$ (see Definition 4.1). Pick sections $s_0, s_1, \ldots, s_r : T \to Y$ with $s_i(t) = p_i$. Let $E = F^* \mathscr{T}_X(-\sum_{i=1}^r s_i(T))$. By Riemann-Roch, for m large enough, we have that $H^1(C, M \otimes E|_C) = 0$ for all line bundles M of degree larger than m, and also that $E|_{C_i}$ is ample since C_i is $(r+1)$-free. Now apply [Kol96, II.7.10.1] for m large enough. ☐

Using any of Theorems 5.3, 5.4 or 5.2, a smooth projective variety X over an algebraically closed field of characteristic zero with a free genus g curve $f : C \to X$ such that $g \geq 1$ is automatically rationally connected.

Remark 5.6. At this point we cannot prove that in positive characteristic, assuming that we have a free curve $f : C \to X$ of genus $g \geq 1$ implies that X is separably rationally connected or even rationally chain connected. It is tempting to hope that both statements are true though. Jason Starr informs us that his maximal free rational quotient (MFRC) [Sta06] gives a generically (on the source) smooth morphism $X \to R_f(X)$ over any algebraically closed field k, so if X contained a free rational curve $f : \mathbb{P}^1 \to X$, then $\dim R_f(X) < \dim X$. Hence, if $f : C \to X$ a free curve of genus $g \geq 1$ implied that we have a free rational curve $\mathbb{P}^1 \to X$ (we do not know how to show this), taking successive MFRC quotients and using Proposition 4.6 would reduce the tower of MFRC quotients to a point. This does not mean that X will necessarily be rationally connected, but since there is a free rational curve on X, it will at least be separably uniruled. Even though Bend and Break arguments give us the existence of many rational curves, the author does not know any general techniques to construct free rational curves in positive characteristic. See the last two sections for results in this direction.

6. Elliptically connected varieties

In this section we will study more carefully the case of genus one. Denoting RC and EC to mean rationally and elliptically connected (genus one connected) respectively, we have the following inclusions of sets of varieties

$$\{\text{rational}\} \subsetneq \{\text{unirational}\} \subseteq \{\text{RC}\} \subsetneq \{\text{EC}\} \subsetneq \{\text{uniruled}\}.$$

It is an open problem whether there exists a non-unirational rationally connected variety but it is widely expected these do exist. The following result is in the spirit of 3.6. The following proof was suggested by the anonymous referee.

PROPOSITION 6.1. *Let X be an elliptically connected smooth projective variety of* $\dim X \geq 2$ *over an algebraically closed field k. Then X is uniruled.*

PROOF. Like in 3.6, for $\mathcal{C} \to U$ a family of genus one curves mapping to X such that $\mathcal{C} \times_U \mathcal{C} \to X \times X$ is dominant, there is an at least one dimensional locus $Z \subset U$ parametrising curves which pass through a (general) point $x \in X$. In fact, after fixing a general hyperplane H, we obtain a morphism $Z \to \mathcal{M}_{1,2}(X)$ where for $z \in Z$, the two marked points are the point $p_z \in \mathcal{C}_z$ sent to x, and a point $q_z \in \mathcal{C}_z$ which is sent to H. Denote also by $\mathcal{C} \to Z$ the restriction of the family from U. Consider now a compactification and the induced rational map to X

$$\begin{tikzcd} \overline{\mathcal{C}} \arrow[r, dashed, "f"] \arrow[d, "\pi"] & X \\ \overline{Z} & \end{tikzcd}$$

and let $\mu : \overline{Z} \to \overline{\mathcal{M}}_{1,2}(X)$ be the moduli map. Since $\mathcal{M}_{1,2}$ contains no proper subvarieties which do not get contracted when mapped to $\mathcal{M}_1$, either the image of μ meets the boundary, which implies that there is a rational curve through x, or μ is a contraction to a point. In the latter case, we thus have that the family π is isotrivial, so after passing to a finite flat cover $\overline{Z}'$ of $\overline{Z}$ we obtain $C \times \overline{Z}' \to \overline{Z}'$, with $f' : C \times \overline{Z}' \dashrightarrow X$ the induced morphism. From the construction, we also obtain a point $p \in C$ (mapped to each p_z under the map $C \times \overline{Z}' \to \overline{\mathcal{C}}$) such that f' contracts $\{p\} \times \overline{Z}'$ to x. If f' were defined everywhere, Mumford's Rigidity Theorem would imply that all fibres $\{s\} \times \overline{Z}'$ are contracted, which contradicts the fact that images of our initial family dominate H. Hence f' is not defined everywhere and like in Bend and Break, we obtain a rational curve through x. $\square$

THEOREM 6.2. *Let X be a smooth projective variety over an algebraically closed field k of characteristic zero. Then X is elliptically connected if and only if it is rationally connected or a rationally connected fibration over an elliptic curve.*

PROOF. Consider the MRC fibration $\pi : X \dashrightarrow R(X)$ where $R(X)$ is elliptically connected as π is dominant. Since $R(X)$ is elliptically connected and not uniruled, it follows from Proposition 6.1 that it must be either of dimension 0 and thus X is rationally connected, or of dimension 1 and so an elliptic curve E by Riemann-Hurwitz. By Remark 3.8, the MRC fibration coincides with the map to the Albanese and so fibres of $X \to E'$ are rationally connected. Conversely, we have seen that a rationally connected variety is elliptically connected in Lemma 3.2. If on the other hand X is a rationally connected fibration over an elliptic curve E then from Proposition 3.5 we know that it is E-connected. □

If k is of positive characteristic, using the same methods as in Lemma 3.7 we deduce that for an elliptically connected variety, the tower of MRC fibrations terminates with a point or a curve.

Remark 6.3. Note that Bjorn Poonen [Poo10] has constructed non-trivial examples over an arbitrary field, of elliptically connected threefolds which are not rationally connected. These are Châtelet surface fibrations over an elliptic curve.

7. Towards a positive characteristic analogue

From Remark 5.6 and the work preceding it, we would like to demonstrate that the existence of a free higher genus curve implies the existence of a free rational curve in positive characteristic, something which holds in characteristic zero from Theorem 5.2. In this section we make the first steps in this direction. If $f : C \to X$ is a very free morphism from a smooth projective curve of genus $g \geq 2$ to a smooth projective variety X, then $K_X.C = -\deg f^*T_X < 0$ from the ampleness of f^*T_X. In fact, a Riemann-Roch calculation gives a better bound of $K_X.C \leq -n(g-1)$ where $n = \dim X$.

Proposition 7.1. *Let X be a smooth projective surface over an algebraically closed field k with $f : C \to X$ a free morphism from a smooth projective curve C of genus $g > 0$ or a very free morphism of genus zero. It follows that X is separably rationally connected.*

Proof. If C is of genus zero then X is separably rationally connected by definition. From the discussion above we have that K_X is not nef. Also, any surface Y which is birational to X admits a morphism $C \to Y$ from 4.6, which is again free, so K_Y is also not nef. From the classification of surfaces this means that X is either rational or ruled. If ruled, X would admit a birational morphism to $\mathbb{P}^1 \times C$. The free morphism $f : C \to X$ would give a free morphism $C \to \mathbb{P}^1 \times C$ which would mean C is $\mathbb{P}^1$ and X was rational. □

Remark 7.2. Some remarks about the case of dimension three, where the minimal model program is incomplete in positive characteristic. From the main theorem in [Kol91], assuming X is smooth and that it admits a free morphism from a curve, we can contract extremal rays in the cone of curves in arbitrary characteristic, to obtain a Fano fibration over a curve, surface or point. In the case where there exists a conic fibration $X \to S$ where S is a smooth surface, Kollár proves that if the characteristic of k is not 2 then the general fibre is smooth. From Proposition 4.6 it follows that the composition morphism $C \to S$ is free and so from the above proposition for the case of surfaces, S is a rational surface. Hence X is a conic bundle over a rational surface hence separably rationally connected. If $X \to Y$ a Fano fibration over a curve, to the author's knowledge, it is not known whether the fibres of the del Pezzo surface fibration over Y obtained in this way must be smooth. Assuming for the time being that they were, they would be separably rationally connected and from the deformation theory argument in Theorem 5.2 and de Franchis' Theorem [ACG11, 8.27], Y would be $\mathbb{P}^1$. From the de Jong-Starr Theorem we would obtain sections $\mathbb{P}^1 \to X$ from which we could assemble combs with very free teeth to be smoothed to very free rational curves in X, showing that X is separably rationally connected. Finally, even though it is open whether Fano threefolds are separably rationally connected (this result is not true in higher dimensions however), Shepherd-Barron [SB97] proved that Fano threefolds of Picard rank one are liftable to characteristic zero, hence admitting a very free morphism implies they are separably rationally connected.

The following result is well known in the case of $\mathbb{P}^1$ (see [Deb01, 4.18]) and easily extends to higher genus.

PROPOSITION 7.3. *Let $f : C \to X$ be a very free morphism from a smooth projective curve C to a smooth projective variety X over an algebraically closed field k. Then for all positive integers m, ℓ*

$$H^0(X, (\Omega^\ell_X)^{\otimes m}) = 0.$$

PROOF. Since $f : C \to X$ is very free, from Proposition 4.7 there is a variety U such that $C \times U \to X$ makes X separably C-connected. Being very free is an open property ([Kol96, II.3.2]) so we can assume that the general morphism $f_u : C_u \to X$ for $u \in U$ is very free and also an immersion from [Kol96, II.1.8], and so $f_u^* \mathscr{T}_X$ is ample from Proposition 2.2 (and by definition of a very free curve in the genus zero case). We conclude that for a general point $x \in X$ there is a morphism $f_u : C_u \to X$ such that $f_u^* \mathscr{T}_X$ is ample and whose image passes through x. Hence since $f_u^* \Omega^1_X$ is negative, any section of $(\Omega^\ell_X)^{\otimes m}$ must vanish on the image $f(C_u)$ hence on a dense open subset of X, and so on X. □

COROLLARY 7.4. *Let $f : C \to X$ as above. Then the Albanese variety* $\mathrm{Alb}\, X$ *is trivial.*

PROOF. Note that we have that $\dim \mathrm{Alb}\, X \leq \dim H^1(X, \mathscr{O}_X) = h^{0,1}$. In characteristic zero Hodge duality gives that $h^{1,0} = h^{0,1}$ but more generally over any algebraically closed field we have (see [Igu55]) that $\dim \mathrm{Alb}\, X \leq h^{1,0} = h^0(X, \Omega^1_X)$. The result follows from Proposition 7.3. □

The above also follows from the result in [Gou14], which says that in the above situation $H^1(X, \mathscr{O}_X) = 0$. See ibid. for a discussion around the vanishing of $H^i(X, \mathscr{O}_X)$ for separably rationally connected varieties in positive characteristic. Note also that if X is C-connected, since any map $C \to \mathrm{Alb}\, X$ must factor through the Jacobian, and there are only countably many homomorphisms between abelian varieties, one concludes that the image of X in $\mathrm{Alb}\, X$ is either a point or a curve.

8. An example in positive characteristic

Let X be the Fermat quintic surface $x_0^5 + x_1^5 + x_2^5 + x_3^5 = 0$ in $\mathbb{P}^3$ over an algebraically closed field of characteristic p. In [Shi74] it is proven that if $p \neq 5$ and p is not congruent to 1 modulo 5, then X is a unirational general type surface and if we quotient by the action of the group G of 5-th roots of unity $x_i \mapsto \zeta^i x_i$, then we obtain a Godeaux surface which is again unirational but has algebraic fundamental group $\pi_1^{\text{et}}(X/G, \overline{y}) \cong \mathbb{Z}/5\mathbb{Z}$. Note that in characteristic zero, the notions of rationally chain connected, rationally connected, freely rationally connected (see [She10]) and separably rationally connected all coincide and it is known that each variety in this class is simply connected. In positive characteristic however these notions are in decreasing generality and can differ. A rationally chain connected variety always has finite fundamental group (see [CL03]) whereas a freely rationally connected variety is simply connected (see [She10]). Note that Shioda's example above gives a unirational and hence rationally connected variety over a characteristic p algebraically closed field which is not simply connected.

We show there is a smooth projective variety in characteristic p which has infinite étale fundamental group but after a finite number of MRC quotients we terminate with a point. Let C be a smooth 5 to 1 cover of $\mathbb{P}^1$, with defining affine equation of the form $y^5 = f(x)$ where f is a general polynomial of high degree. We have an action of $G = \mathbb{Z}/5\mathbb{Z}$ on C which we can extend to the product $X \times C$ of the above Fermat quintic X with C. Projecting from the quotient onto the second factor we have a morphism $(X \times C)/G \to \mathbb{P}^1$ where we have identified C/G with $\mathbb{P}^1$. The general fibre of this morphism is isomorphic to X. We have a short exact sequence

$$1 \to \pi_1^{\text{et}}(X, \overline{x}) \times \pi_1^{\text{et}}(C, \overline{c}) \to \pi_1^{\text{et}}((X \times C)/G, \overline{z}) \to G \to 1.$$

Hence we have constructed an example of a smooth projective variety over an algebraically closed field of characteristic p whose fundamental group is infinite yet whose tower of MRC quotients terminates with a point.

References

[ACG11] Enrico Arbarello, Maurizio Cornalba, and Pillip A. Griffiths. *Geometry of algebraic curves. Volume II*, volume 268 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Heidelberg, 2011. With a contribution by Joseph Daniel Harris.

[AK03] Carolina Araujo and János Kollár. Rational curves on varieties. In *Higher dimensional varieties and rational points (Budapest, 2001)*, volume 12 of *Bolyai Soc. Math. Stud.*, pages 13–68. Springer, Berlin, 2003.

[BDPP13] Sébastien Boucksom, Jean-Pierre Demailly, Mihai Păun, and Thomas Peternell. The pseudo-effective cone of a compact Kähler manifold and varieties of negative Kodaira dimension. *J. Algebraic Geom.*, 22(2):201–248, 2013.

[BM01] Fedor Bogomolov and Michael McQuillan. Rational curves on foliated varieties. *preprint*, 2001.

[Cam92] F. Campana. Connexité rationnelle des variétés de Fano. *Ann. Sci. École Norm. Sup. (4)*, 25(5):539–545, 1992.

[CL03] Antoine Chambert-Loir. A propos du groupe fondamental des variétés rationnellement connexes. 2003.

[Deb01] Olivier Debarre. *Higher-dimensional algebraic geometry*. Universitext. Springer-Verlag, New York, 2001.

[DG67] Jean Dieudonné and Alexander Grothendieck. Éléments de géométrie algébrique. *Inst. Hautes Études Sci. Publ. Math.*, 4, 8, 11, 17, 20, 24, 28, 32, 1961–1967.

[dJS03] A. J. de Jong and J. Starr. Every rationally connected variety over the function field of a curve has a rational point. *Amer. J. Math.*, 125(3):567–580, 2003.

[FP97] W. Fulton and R. Pandharipande. Notes on stable maps and quantum cohomology. In *Algebraic geometry—Santa Cruz 1995*, volume 62 of *Proc. Sympos. Pure Math.*, pages 45–96. Amer. Math. Soc., Providence, RI, 1997.

[GHS03] Tom Graber, Joe Harris, and Jason Starr. Families of rationally connected varieties. *J. Amer. Math. Soc.*, 16(1):57–67 (electronic), 2003.

[Gou14] Frank Gounelas. The first cohomology of separably rationally connected varieties. *C. R. Math. Acad. Sci. Paris*, 352(11):871–873, 2014.

[Har66] Robin Hartshorne. Ample vector bundles. *Inst. Hautes Études Sci. Publ. Math.*, (29):63–94, 1966.

[Har70] Robin Hartshorne. *Ample subvarieties of algebraic varieties*. Notes written in collaboration with C. Musili. Lecture Notes in Mathematics, Vol. 156. Springer-Verlag, Berlin, 1970.

[Har71] Robin Hartshorne. Ample vector bundles on curves. *Nagoya Math. J.*, 43:73–89, 1971.

[Har77] Robin Hartshorne. *Algebraic geometry*. Springer-Verlag, New York, 1977. Graduate Texts in Mathematics, No. 52.

[Igu55] Jun-ichi Igusa. A fundamental inequality in the theory of Picard varieties. *Proc. Nat. Acad. Sci. U.S.A.*, 41:317–320, 1955.

[KMM92] János Kollár, Yoichi Miyaoka, and Shigefumi Mori. Rationally connected varieties. *J. Algebraic Geom.*, 1(3):429–448, 1992.

[Kol91] János Kollár. Extremal rays on smooth threefolds. *Ann. Sci. École Norm. Sup. (4)*, 24(3):339–361, 1991.

[Kol96] János Kollár. *Rational curves on algebraic varieties*, volume 32 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]*. Springer-Verlag, Berlin, 1996.

[KSCT07] Stefan Kebekus, Luis Solá Conde, and Matei Toma. Rationally connected foliations after Bogomolov and McQuillan. *J. Algebraic Geom.*, 16(1):65–81, 2007.

[Laz04] Robert Lazarsfeld. *Positivity in algebraic geometry. II*, volume 49 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]*. Springer-Verlag, Berlin, 2004. Positivity for vector bundles, and multiplier ideals.

[Liu02] Qing Liu. *Algebraic geometry and arithmetic curves*, volume 6 of *Oxford Graduate Texts in Mathematics*. Oxford University Press, Oxford, 2002. Translated from the French by Reinie Erné, Oxford Science Publications.

[Miy87] Yoichi Miyaoka. Deformations of a morphism along a foliation and applications. In *Algebraic geometry, Bowdoin, 1985 (Brunswick, Maine, 1985)*, volume 46 of *Proc. Sympos. Pure Math.*, pages 245–268. Amer. Math. Soc., Providence, RI, 1987.

[Mor79] Shigefumi Mori. Projective manifolds with ample tangent bundles. *Ann. of Math. (2)*, 110(3):593–606, 1979.

[Occ06] Gianluca Occhetta. Extending rationally connected fibrations. *Forum Math.*, 18(5):853–867, 2006.

[Ott12] John Christian Ottem. Ample subvarieties and q-ample divisors. *Adv. Math.*, 229(5):2868–2887, 2012.

[Pet06] Thomas Peternell. Kodaira dimension of subvarieties. II. *Internat. J. Math.*, 17(5):619–631, 2006.

[Poo10] Bjorn Poonen. Insufficiency of the Brauer-Manin obstruction applied to étale covers. *Ann. of Math. (2)*, 171(3):2157–2169, 2010.

[SB97] N. I. Shepherd-Barron. Fano threefolds in positive characteristic. *Compositio Math.*, 105(3):237–265, 1997.

[She10] Mingmin Shen. Foliations and rational connectedness in positive characteristic. *J. Algebraic Geom.*, 19(3):531–553, 2010.

[Shi74] Tetsuji Shioda. An example of unirational surfaces in characteristic p. *Math. Ann.*, 211:233–236, 1974.

[Sta06] Jason Starr. The maximal free rational quotient. *preprint*, 2006.

Frank Gounelas
Institut für Mathematik
Humboldt Universität zu Berlin
Unter den Linden 6
10099 Berlin
gounelas@mathematik.hu-berlin.de

The Isomorphism Problem for Semigroup C^*-Algebras of Right-Angled Artin Monoids

Søren Eilers, Xin Li, and Efren Ruiz

Received: February 2, 2016

Communicated by Joachim Cuntz

Abstract. Semigroup C*-algebras for right-angled Artin monoids were introduced and studied by Crisp and Laca. In the paper at hand, we are able to present the complete answer to their question of when such C*-algebras are isomorphic. The answer to this question is presented both in terms of properties of the graph defining the Artin monoids as well as in terms of classification by K-theory, and is obtained using recent results from classification of non-simple C*-algebras. Moreover, we are able to answer another natural question: Which of these semigroup C*-algebras for right-angled Artin monoids are isomorphic to graph algebras? We give a complete answer, and note the consequence that many of the C*-algebras under study are semiprojective.

2010 Mathematics Subject Classification: Primary: 46L05. Secondary: 20F36, 20M05, 46L35.

1 Introduction

Semigroup C*-algebras for right-angled Artin monoids were introduced and studied by Crisp and Laca in [CL02] and [CL07]. In [CL07], the authors ask how to classify these semigroup C*-algebras up to *-isomorphism. We now present the complete answer to their question.

The Artin monoids studied here are given by countable, symmetric and antireflexive graphs $\Gamma = (V, E)$ as

$$A_\Gamma^+ := \langle \{\sigma_v \colon v \in V\} \,|\sigma_v \sigma_w = \sigma_w \sigma_v \text{ if } (v, w) \in E \rangle^+ .$$

The corresponding right-angled Artin groups, defined by the same generators and relations, are special cases of Artin groups, which form an important class of examples of groups. We refer the reader to [CL02, CL07] and the references therein for more details.

Semigroup C*-algebras of left cancellative semigroups, generated by the left regular representation of the semigroup, have been studied for a long time. Recently, there has been a renewed interest in this topic (see [Li12, Li13] and the references therein). By [CL02], the semigroup C*-algebras $C^*(A_\Gamma^+)$ attached to right-angled Artin monoids are given as the universal C*-algebras for

$$\left\langle \{s_v\colon v \in V\} \;\middle|\; \begin{array}{c} [s_v, s_w] = [s_v, s_w^*] = 0 \text{ if } (v,w) \in E \\ s_v^* s_w = \delta_{v,w} \text{ if } (v,w) \notin E \end{array} \right\rangle$$

We answer the question of when two graphs Γ, Λ produce C*-algebras that are isomorphic. Although we emphasize that our results cover the full range of such graphs, it is instructive to state our main results in the case of finite graphs. This is a specialization of the combination of Theorems 4.2 and 5.2.

Theorem 1.1 *Let Γ and Λ be finite undirected graphs with no loops. The following are equivalent*

1. $C^*(A_\Gamma^+) \cong C^*(A_\Lambda^+)$
2. (a) $t(\Gamma) = t(\Lambda)$
   (b) $N_k(\Gamma) + N_{-k}(\Gamma) = N_k(\Lambda) + N_{-k}(\Lambda)$ *for all* $k \in \mathbb{Z}$
   (c) $N_0(\Gamma) > 0$ *or*
$$\sum_{k>0} N_{-k}(\Gamma) \equiv \sum_{k>0} N_{-k}(\Lambda) \mod 2$$
3. $[FK_+(C^*(A_\Gamma^+)), [1_{C^*(A_\Gamma^+)}]] \cong [FK_+(C^*(A_\Lambda^+)), [1_{C^*(A_\Lambda^+)}]]$

In this result, the invariant mentioned in (3) is the standard ordered filtered K-theory – implicitly containing the primitive ideal space – which has been conjectured in [ERR10] to be a complete invariant for a large and important class of C*-algebras. This conjecture is still open, but has been confirmed in a multitude of situations partially overlapping with the case at hand. But the main strength of our result is in fact that the *ad hoc* invariant of (2) is extremely easy to compute for Γ and Λ. Indeed, as we shall detail below, the numbers $t(\Gamma)$ and $N_k(\Gamma)$ are obtained by dividing Γ into co-irreducible components and then counting how many of these are singletons, yielding $t(\Gamma)$, and counting how many of the remaining co-irreducible components have Euler characteristic k, yielding $N_k(\Gamma)$. In Figure 1 this process has been completed for all 34 graphs with five vertices, and we conclude that they define 18 different C*-algebras.

When the number of vertices increase, it is possible for two graphs to have different sets of invariants, yet define the same C*-algebra. For instance, defining

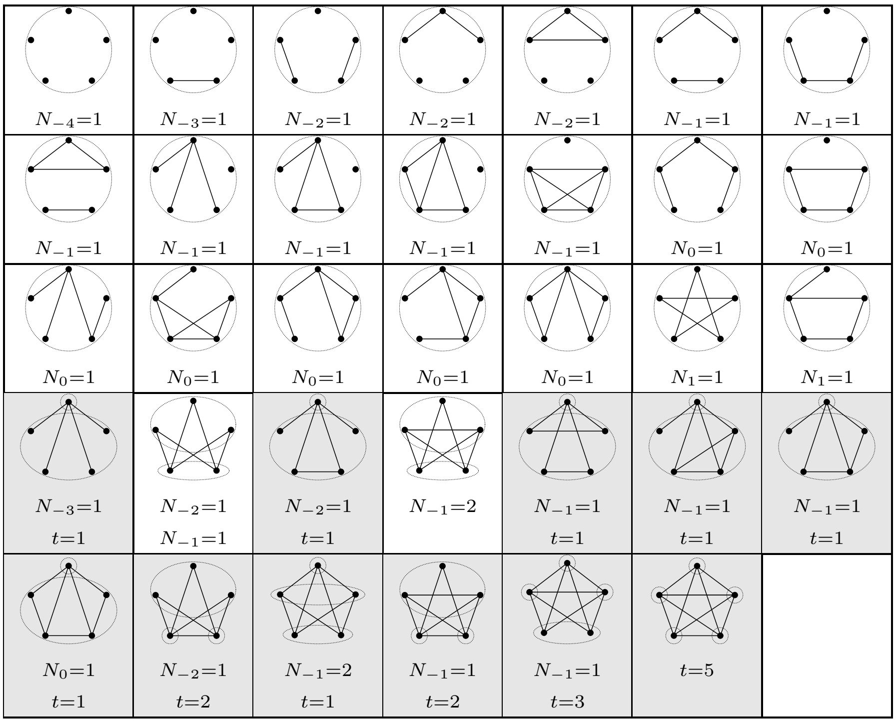


Figure 1: Invariants for all graphs with 5 vertices. Any quantity not mentioned equals zero.

a graph Γ' with 10 vertices having its co-irreducible components chosen among those given in Figure 1 so that

$$N_{-1}(\Gamma') = 2$$

and Λ' similarly defined so that

$$N_1(\Lambda') = 2,$$

then $C^*(A^+_{\Gamma'})$ and $C^*(A^+_{\Lambda'})$ will be isomorphic. Similarly, we may define Γ'' and Λ'' with 15 vertices each so that

$$N_{-1}(\Gamma'') = 1, N_0(\Gamma'') = 1, N_1(\Gamma'') = 1$$
$$N_{-1}(\Lambda'') = 2, N_0(\Lambda'') = 1$$

obtaining that $C^*(A^+_{\Gamma''}) \cong C^*(A^+_{\Lambda''})$.

In the general case of possibly infinite graphs, an additional quantity $o(\Gamma)$ must be introduced to count those co-irreducible components which have an infinite number of vertices, and to address the possibility of having an infinite number of co-irreducible components, but the necessary condition replacing (2) in this general case is not much more complicated than the one given above.

We note that the C*-algebras associated via semigroups to undirected and loop-free graphs are not always graph C*-algebras in the usual sense, not only because graph C*-algebras are defined using directed graphs. We provide a complete description of when $C^*(A^+_\Gamma)$ is in fact a graph C*-algebra, and note that there is a rather complicated relation between Γ and the G_Γ when in fact $C^*(A^+_\Gamma) \cong C^*(G_\Gamma)$. In this case, G_Γ is not unique.

Our results have surprising consequences for the issue of *stable relations* (cf. [Lor97]) among sets of isometries of separable Hilbert spaces, subject to commutativity or orthogonality relations as given by the graph Γ, or, which is nearly the same, for the issue of *semiprojectivity* (cf. [Bla85]) of the C*-algebras $C^*(A^+_\Gamma)$. Indeed, it is easy to see by spectral theory that $C^*(A^+_\Gamma)$ is semiprojective when Γ is a finite graph with no edges, corresponding to a family of isometries having orthogonal range projections. Similarly, it follows e.g. from considering the celebrated Voiculescu matrices ([Voi83],[EL91]) that when Γ is a complete graph with more than one vertex, $C^*(A^+_\Gamma)$ cannot have this property. It is a question attracting a lot of attention (see e.g. [Bla04]) to what extent it is possible to obtain stable relations for commuting sets of stable relations, or to what extent tensor products of semiprojective C*-algebras can themselves be semiprojective. In fact, it was only established recently ([End15]) that $\mathcal{O}_3 \otimes \mathcal{O}_3$ is semiprojective, as a consequence of Enders' sweeping solution to the semiprojectivity problem for Kirchberg algebras. In our setting, because we have found that many settings in which some isometries are required to be orthogonal, and others to commute, give the same C*-algebras as the ones where all are required to be orthogonal, we immediately see that many such settings — for instance the first 12 listed in Figure 1 — provide for stable relations. Involving the notion of graph algebras as outlined above, we will show in Theorem 6.9 semiprojectivity and nonsemiprojectivity for many of the C*-algebras under study, and it follows from our results that exactly those C*-algebras arising from the graphs in Figure 1 in the non-shaded entries are semiprojective. We have not been able to resolve the issue completely as Enders' methods do not apply directly, the first open case having six vertices and two co-irreducible components each with Euler characteristic -2.

All authors acknowledge support by the Danish National Research Foundation through the Centre for Symmetry and Deformation (DNRF92), and thank the Department of Mathematical Sciences at the University of Copenhagen, where the initial phases of this work were carried out, for providing excellent facilities. The first named author gratefully acknowledges support from the VILLUM Foundation, and the third named author gratefully acknowledges support from the Simons Foundation (#279369 to Efren Ruiz).

2 Preliminaries

2.1 Semigroup C*-algebras for right-angled Artin monoids

Let Γ be a countable graph. $\Gamma = (V, E)$ is given by a countable set of vertices V and a set of edges E. We only consider unoriented edges, and given two vertices, there is at most one edge joining these two vertices. In other words, we can think of E as a symmetric subset of $V \times V$ not containing elements of the diagonal.

Given such a graph $\Gamma = (V, E)$, the right-angled Artin group A_Γ is defined as follows:

$$A_\Gamma := \langle \{\sigma_v \colon v \in V\} \,|\sigma_v\sigma_w = \sigma_w\sigma_v \text{ if } (v,w) \in E\rangle \,.$$

Similarly, the right-angled Artin monoid A_Γ^+ is defined as follows:

$$A_\Gamma^+ := \langle \{\sigma_v \colon v \in V\} \,|\sigma_v\sigma_w = \sigma_w\sigma_v \text{ if } (v,w) \in E\rangle^+ \,.$$

It turns out that the canonical semigroup homomorphism $A_\Gamma^+ \to A_\Gamma$ is injective, see [Par02]. Moreover, it is shown in [CL02] that $A_\Gamma^+ \subseteq A_\Gamma$ is quasi-lattice ordered. This means that for every g in A_Γ, either $(gA_\Gamma^+) \cap A_\Gamma^+ = \emptyset$ or there exists $p \in A_\Gamma^+$ with $(gA_\Gamma^+) \cap A_\Gamma^+ = pA_\Gamma^+$.

The (left) reduced semigroup C*-algebra of A_Γ^+ is given by

$$C^*_\lambda(A_\Gamma^+) = C^* \langle \{S_v \colon v \in V\}\rangle \subseteq \mathcal{L}(\ell^2(A_\Gamma^+)),$$

where $S_v e_x = e_{\sigma_v x}$ with $\{e_x\}$ the canonical orthonormal basis. The full semigroup C*-algebra of A_Γ^+ is defined as

$$C^*(A_\Gamma^+) = C^*_{\text{univ}} \left\langle \{s_v \colon v \in V\} \,\middle|\, \begin{array}{c} [s_v, s_w] = [s_v, s_w^*] = 0 \text{ if } (v,w) \in E \\ s_v^* s_w = \delta_{v,w} \text{ if } (v,w) \notin E \end{array} \right\rangle$$

The canonical homomorphism $C^*(A_\Gamma^+) \to C^*_\lambda(A_\Gamma^+)$ is an isomorphism by [CL02]. Hence we do not distinguish between reduced and full versions and simply write $C^*(A_\Gamma^+)$ for the semigroup C*-algebra of A_Γ^+.

2.2 Co-irreducible components

The graph Γ is called co-reducible if there exist non-empty subsets V_1 and V_2 of V with $V = V_1 \sqcup V_2$ such that $V_1 \times V_2 \subseteq E$. Γ is called co-irreducible if Γ is not co-reducible. In general, we can always decompose Γ into co-irreducible components. This means that there exist co-irreducible graphs $\Gamma_i = (V_i, E_i)$ such that $A_\Gamma^+ \cong \bigoplus_i A_{\Gamma_i}^+$ (and also $A_\Gamma \cong \bigoplus_i A_{\Gamma_i}$). As explained in [CL07], these co-irreducible components are found by looking at the opposite graph of Γ. For the semigroup C*-algebra, we get $C^*(A_\Gamma^+) \cong \bigotimes_i C^*(A_{\Gamma_i}^+)$. Note that if there are (necessarily countably) infinitely many co-irreducible components, the tensor product is defined as an inductive limit of finite tensor products with respect to the canonical unital embeddings as tensor factors.
It is shown in [CL07] that for a co-irreducible graph $\Gamma = (V, E)$ with $1 < |V| < \infty$, $C^*(A_\Gamma^+)$ has a unique non-trivial ideal isomorphic to the compact operators. It is easy to see the compact operators in the description of $C^*(A_\Gamma^+)$ as a concrete C*-algebra on $\ell^2(A_\Gamma^+)$: We just have to observe that $1 - \bigvee_{v \in V} S_v S_v^*$ is the orthogonal projection onto the one-dimensional subspace of $\ell^2(A_\Gamma^+)$ corresponding to the identity element of A_Γ^+. This projection then generates the ideal of compact operators. The corresponding quotient $C_Q^*(A_\Gamma^+)$ is a (unital) Kirchberg algebra satisfying the UCT. However, if our co-irreducible graph has infinitely many vertices, then $C^*(A_\Gamma^+)$ itself is a (unital) Kirchberg algebra satisfying the UCT. That we obtain UCT Kirchberg algebras follows also from [Li13, Corollary 7.23]. The case where Γ consists of only one vertex is easy to understand; in that case, $C^*(A_\Gamma^+)$ is canonically isomorphic to the Toeplitz algebra $\mathcal{T}$.

2.3 Primitive ideal space

We can now describe the primitive ideal space of $C^*(A_\Gamma^+)$ for arbitrary Γ. Let $\Gamma_i = (V_i, E_i)$ be the co-irreducible components of Γ. Then by [Bla77, Theorem 4.9], we have an identification

$$\mathrm{Prim}\,(C^*(A_\Gamma^+)) \cong \prod_i \mathrm{Prim}\,(C^*(A_{\Gamma_i}^+)).$$

Under this identification, an element (I_i) of the space on the right hand side corresponds to the primitive ideal I of $C^*(A_\Gamma^+)$ which is generated by $\{\bigotimes_j J_{ij}\}_i$, where $J_{ij} = C^*(A_{\Gamma_j}^+)$ if $j \neq i$ and $J_{ii} = I_i$. Since each of the Γ_i is co-irreducible, the primitive ideal space $\mathrm{Prim}\,(C^*(A_{\Gamma_i}^+))$ is easy to describe because of the results summarized above:

- If Γ_i just consists of one point, then $\mathrm{Prim}\,(C^*(A_{\Gamma_i}^+))$ is homeomorphic to the primitive ideal space of the Toeplitz algebra. This means that as a set, $\mathrm{Prim}\,(C^*(A_{\Gamma_i}^+))$ is the disjoint union of a point and a circle, and the non-empty open sets are given by unions of the point and open subsets in the usual topology of the circle.

- If Γ_i has more than one, but finitely many vertices, $\mathrm{Prim}\,(C^*(A_{\Gamma_i}^+))$ consists of two points, one of which is closed (the corresponding primitive ideal is the ideal of compact operators) and the other one is dense.
- If Γ_i has infinitely many vertices, then $\mathrm{Prim}\,(C^*(A_{\Gamma_i}^+))$ consists of only one point.

2.4 K-theory

K-theory for $C^*(A_\Gamma^+)$ and the quotients $C_Q^*(A_\Gamma^+)$ has been computed in [Iva10] in an ad hoc way, and can also be computed using [CEL13]. Let us explain the computation via the latter route. First of all, we need the Euler characteristic of a graph Γ. We view Γ as a simplicial complex by defining for every $n = 0, 1, 2, \ldots$ the set of n-simplices by

$$K_n := \{\{v_0, \ldots, v_n\} \subseteq V\colon\ (v_i, v_j) \in E \text{ for all } i, j \in \{0, \ldots, n\},\, i \neq j\}\,.$$

Then we set for a graph Γ with finitely many vertices $\chi(\Gamma) := 1 - \sum_{n=0}^{\infty} (-1)^n |K_n|$.

REMARK 2.1 *It is easy to see that there are co-irreducible graphs attaining any integer as its Euler characteristic. Indeed, letting Γ_{-m} denote the graph with $m+1$ vertices and no edges, we clearly have*

$$\chi(\Gamma_{-m}) = -m.$$

Systematically generating positive characteristics is harder; one option is to let Γ_{n^2-1} denote the graph with $2n+2$ vertices obtained by deleting one edge from the complete bipartite graph $K_{n+1,n+1}$ and note that

$$\chi(\Gamma_{n^2-1}) = n^2 - 1$$

To obtain positive characteristics in $\{(n-1)^2, \ldots, n^2-1\}$ one may simply add a suitable number of isolated vertices to Γ_{n^2-1}.

Now, by [CEL13, Theorem 5.2], we know that we always have $K_*(C^*(A_\Gamma^+)) \cong K_*(\mathbb{C})$, and $K_0(C^*(A_\Gamma^+)) \cong \mathbb{Z}$ is generated by the class of the unit $[1]$. Here we use that right-angled Artin groups satisfy the Baum-Connes conjecture with coefficients because these groups have the Haagerup property (see [NR97], and also [AD]). To compute K-theory for the quotient $C_Q^*(A_\Gamma^+)$ in the case that Γ has (more than one and) finitely many vertices, we consider the short exact sequence $0 \to \mathcal{K} \to C^*(A_\Gamma^+) \to C_Q^*(A_\Gamma^+) \to 0$ and its six term exact sequence in K-theory:

$$\begin{tikzcd}
K_0(\mathcal{K}) \arrow[r] & K_0(C^*(A_\Gamma^+)) \arrow[r] & K_0(C_Q^*(A_\Gamma^+)) \arrow[d] \\
K_1(C_Q^*(A_\Gamma^+)) \arrow[u] & K_1(C^*(A_\Gamma^+)) \arrow[l] & K_1(\mathcal{K}) \arrow[l]
\end{tikzcd} \tag{1}$$

Since both $K_1(\mathcal{K})$ and $K_1(C^*(A_\Gamma^+))$ vanish, all we have to do is to compute the homomorphism $K_0(\mathcal{K}) \cong \mathbb{Z} \to \mathbb{Z} \cong K_0(C^*(A_\Gamma^+))$. $K_0(\mathcal{K}) \cong \mathbb{Z}$ is generated by the class of any minimal projection. So we can take $e = 1 - \bigvee_{v \in V} S_v S_v^*$. It is easy to see that in K_0, $[e] \in K_0(\mathcal{K})$ is sent to $\chi(\Gamma)[1] \in K_0(C^*(A_\Gamma^+))$. Therefore, by exactness of (1), we conclude that $K_0(C_Q^*(A_\Gamma^+)) \cong \mathbb{Z}/|\chi(\Gamma)|\mathbb{Z}$ and

$$K_1(C_Q^*(A_\Gamma^+)) \cong \begin{cases} \{0\} & \text{if } \chi(\Gamma) \neq 0 \\ \mathbb{Z} & \text{if } \chi(\Gamma) = 0 \end{cases} .$$

3 Extension algebras

We now discuss the C*-algebras associated to co-irreducible graphs and see how they are all isomorphic to either the Toeplitz algebra, the Cuntz algebra $\mathcal{O}_\infty$, or an *extension algebra* as specified below.

Theorem 3.1 *Consider the following properties for a unital C*-algebra A:*

(1) A contains $\mathcal{K}$ as an ideal, and $A/\mathcal{K}$ is a Kirchberg algebra satisfying the UCT,

(2) $K_0(A) = \mathbb{Z}$ with $[1_A] = 1$.

For each $k \in \mathbb{Z}\backslash\{0\}$ there is a unique C-algebra satisfying (1), (2) and*

(3) The six-term exact sequence for $\mathcal{K}$ and A is given by

$$\begin{tikzcd} \mathbb{Z} \arrow[r, "k"] & \mathbb{Z} \arrow[r] & \mathbb{Z}/k\mathbb{Z} \arrow[d] \\ 0 \arrow[u] & 0 \arrow[l] & 0 \arrow[l] \end{tikzcd}$$

There is also a unique C-algebra satisfying (1), (2) and*

(3') The six-term exact sequence for $\mathcal{K}$ and A is given by

$$\begin{tikzcd} \mathbb{Z} \arrow[r, "0"] & \mathbb{Z} \arrow[r, equal] & \mathbb{Z} \arrow[d] \\ \mathbb{Z} \arrow[u, equal] & 0 \arrow[l] & 0 \arrow[l] \end{tikzcd}$$

Proof: Note that $\mathcal{K}$ is an essential ideal of A (i.e., every nonzero ideal of A has a nontrivial intersection with $\mathcal{K}$) since A is unital and $A/\mathcal{K}$ is simple. Uniqueness follows from [ERR, Corollary 4.22]. For existence, we note that when Γ is a finite and co-irreducible graph with $|\Gamma| > 1$ and $\chi(\Gamma) = k$, all properties are met as noted in Section 2. □

When specifying the map $K_0(\mathcal{K}) \to K_0(A)$ above we let the unit of the leftmost copy of $\mathbb{Z}$ denote the class of a minimal projection of $\mathbb{Z}$.

Definition 3.2 *The unique C^*-algebras satisfying (1),(2) and (3) are denoted $E_{|k|+1}^{\mathrm{sgn}(k)}$. The unique C^*-algebra satisfying (1),(2) and (3') is denoted E_1^0. The quotient $E_1^0/\mathcal{K}$ is denoted $\mathcal{O}_1$.*

Our notation has been chosen to fit the notation E_n^k for the extension algebras of $\mathcal{O}_n$ studied in [FL07]. With our name $\mathcal{O}_1$ for the appropriately chosen Kirchberg algebra, we have

$$0 \longrightarrow \mathcal{K} \xrightarrow{\iota} E_n^k \xrightarrow{\pi} \mathcal{O}_n \longrightarrow 0$$

for any $k \in \{-1, 0, 1\}$ and $n \in \mathbb{N}$, provided $k = 0$ precisely when $n = 1$.

Lemma 3.3 *$E_n^k \cong E_{n'}^{k'}$ only when $n = n'$ and $k = k'$. $E_n^k \otimes \mathcal{K} \cong E_{n'}^{k'} \otimes \mathcal{K}$ precisely when $n = n'$.*

Proof: Since the six-term exact sequences are as specified in (3) or (3') of Theorem 3.1, stable isomorphism can only occur if $n = n'$, and hence we only need to check that for $n > 0$, we have $E_n^1 \ncong E_n^{-1}$, yet $E_n^1 \otimes \mathcal{K} \cong E_n^{-1} \otimes \mathcal{K}$.
We note that the only two options for an isomorphism among the six-term exact sequences in this case are given as

$$\begin{array}{ccccc}
\mathbb{Z} & \xrightarrow{n} & \mathbb{Z} & \longrightarrow & \mathbb{Z}/n\mathbb{Z} \\
\downarrow{\scriptstyle \pm 1} & & \downarrow{\scriptstyle \mp 1} & & \downarrow \\
\mathbb{Z} & \xrightarrow[-n]{} & \mathbb{Z} & \longrightarrow & \mathbb{Z}/n\mathbb{Z}
\end{array}$$

and that we must choose $+1$ as the left most isomorphism to preserve the positive cone of $K_0(\mathcal{K})$. Thus, an isomorphism is ruled out as it would fail to send the class of the unit of E_n^1 to the unit of E_n^{-1}, but an isomorphism after stabilization is guaranteed by, e.g., [ERR09]. $\square$
The following result follows directly from § 2.2, § 2.4 and Theorem 3.1.

Theorem 3.4 *When Γ is a co-irreducible graph, $C^*(A_\Gamma^+)$ is one of the C^*-algebras $\mathcal{T}, E_n^k, \mathcal{O}_\infty$ according to*

1. *If $|\Gamma| = 1$, $C^*(A_\Gamma^+) \cong \mathcal{T}$*
2. *If $1 < |\Gamma| < \infty$, $C^*(A_\Gamma^+) \cong E_{1+|\chi(\Gamma)|}^{\mathrm{sgn}\,\chi(\Gamma)}$*
3. *If $|\Gamma| = \infty$, $C^*(A_\Gamma^+) \cong \mathcal{O}_\infty$*

We note that by the information already noted on the ideal structures in combination with Lemma 3.3, the C*-algebras appearing are not mutually isomorphic, and hence we have a complete classification by the cardinality of Γ and the Euler characteristic in the co-irreducible case.
In preparation for the general case we now study isomorphisms between various tensor products amongst the relevant extension algebras and some of their quotients. For this, we will need:

Theorem 3.5 *Let A_i, $i = 1, 2$, be unital C*-algebras whose proper ideals are precisely given by (0), I_i, J_i and $I_i \oplus J_i$. We assume that I_i and J_i are UCT Kirchberg algebras, and the quotients $A_i/(I_i \oplus J_i)$ are also UCT Kirchberg algebras.*
Let $\alpha_I : K_*(I_1) \cong K_*(I_2)$, $\alpha_J : K_*(J_1) \cong K_*(J_2)$, $\alpha_{I\oplus J} : K_*(I_1 \oplus J_1) \cong K_*(I_2 \oplus J_2)$, $\beta : K_*(A_1) \cong K_*(A_2)$, $\gamma_I : K_*(A/I_1) \cong K_*(A/I_2)$, $\gamma_J : K_*(A/J_1) \cong K_*(A/J_2)$, and $\gamma_{I\oplus J} : K_*(A_1/(I_1 \oplus J_1)) \cong K_*(A_2/(I_2 \oplus J_2))$ be isomorphisms, with β preserving the K_0-classes of the units and $\alpha_{I\oplus J} = \alpha_I \oplus \alpha_J$ (under the canonical isomorphism $K_*(I_i \oplus J_i) \cong K_*(I_i) \oplus K_*(J_i)$).
Furthermore, we assume that these isomorphism are compatible with the K-theoretic six term exact sequences attached to

$$0 \to I_i \to A_i \to A_i/I_i \to 0, \quad 0 \to J_i \to A_i \to A_i/J_i \to 0,$$
$$0 \to I_i \oplus J_i \to A_i \to A_i/(I_i \oplus J_i) \to 0, \quad 0 \to J_i \to A_i/I_i \to A_i/(I_i \oplus J_i) \to 0$$

and

$$0 \to I_i \to A_i/J_i \to A_i/(I_i \oplus J_i) \to 0.$$

Then there exists an isomorphism $\varphi : A_1 \cong A_2$ which induces α_I, α_J, $\alpha_{I\oplus J}$, β, γ_I, γ_J and $\gamma_{I\oplus J}$ in K-theory.

Proof: Combine [Kir00, Folgerung 4.3] and [BK, Theorem 1.3] with [RR07, Theorem 2.1] or [ERR, Theorem 3.3]. □

Lemma 3.6 *For every $n \geq 2$, we have $\mathcal{O}_\infty \otimes E_n^{+1} \cong \mathcal{O}_\infty \otimes E_n^{-1}$.*

Proof: Both $\mathcal{O}_\infty \otimes E_n^{+1}$ and $\mathcal{O}_\infty \otimes E_n^{-1}$ are unital C*-algebras with unique ideal isomorphic to $\mathcal{O}_\infty \otimes \mathcal{K}$ and corresponding quotient isomorphic to $\mathcal{O}_\infty \otimes \mathcal{O}_n \cong \mathcal{O}_n$. The K-theoretic six term exact sequences for $0 \to \mathcal{O}_\infty \otimes \mathcal{K} \to \mathcal{O}_\infty \otimes E_n^{+1} \to \mathcal{O}_n \to 0$ and $0 \to \mathcal{O}_\infty \otimes \mathcal{K} \to \mathcal{O}_\infty \otimes E_n^{-1} \to \mathcal{O}_n \to 0$ look as follows:

$$\begin{tikzcd} \mathbb{Z} \arrow[r] & \mathbb{Z} \arrow[r] & \mathbb{Z}/(n-1)\mathbb{Z} \arrow[d] \\ 0 \arrow[u] & 0 \arrow[l] & 0 \arrow[l] \end{tikzcd}$$

where $\mathbb{Z} \cong K_0(\mathcal{O}_\infty \otimes \mathcal{K})$ is generated by $[1 \otimes e]$ for a minimal projection $e \in \mathcal{K}$ and the unit 1 of $\mathcal{O}_\infty$, and $\mathbb{Z} \cong K_0(\mathcal{O}_\infty \otimes E_n^{\pm 1})$ is generated by the class of the unit. The only difference is that for E_n^{+1}, the homomorphism $\mathbb{Z} \to \mathbb{Z}$ is given by $z[1 \otimes e] \mapsto (n-1)z[1]$, whereas for E_n^{-1}, the homomorphism $\mathbb{Z} \to \mathbb{Z}$ is given by $z[1 \otimes e] \mapsto -(n-1)z[1]$ (for $z \in \mathbb{Z}$). We now apply [RR07, Theorem 2.2] to $I_i = \mathcal{O}_\infty \otimes \mathcal{K}$, $A_1 = \mathcal{O}_\infty \otimes E_n^{+1}$, $A_2 = \mathcal{O}_\infty \otimes E_n^{-1}$, $Q_i = \mathcal{O}_n$ and the homomorphisms $\alpha = -\mathrm{id}_{K_0(\mathcal{O}_\infty \otimes \mathcal{K})}$, $\beta : K_0(\mathcal{O}_\infty \otimes E_n^{+1}) \to K_0(\mathcal{O}_\infty \otimes E_n^{-1})$, $z[1] \mapsto z[1]$ (for $z \in \mathbb{Z}$), $\gamma = \mathrm{id}_{K_0(\mathcal{O}_n)}$. It is then obvious that all the assumptions in [RR07] are satisfied, and we conclude that $\mathcal{O}_\infty \otimes E_n^{+1} \cong \mathcal{O}_\infty \otimes E_n^{-1}$. □

Now recall that we have introduced the extension algebra E_1^0 in Theorem 3.1. The C*-algebra $E_1^0 \otimes E_n^{+1}$ $(n \geq 2)$ contains the ideal $\mathcal{K} \otimes \mathcal{K} \cong \mathcal{K}$, and we denote the corresponding quotient by Q^+. Obviously, the primitive ideals of Q^+ are given by $\mathcal{K} \otimes \mathcal{O}_n$, $\mathcal{O}_1 \otimes \mathcal{K}$ and $\mathcal{K} \otimes \mathcal{O}_n \oplus \mathcal{O}_1 \otimes \mathcal{K}$. From the six term exact sequence in K-theory for $0 \to \mathcal{K} \to E_1^0 \otimes E_n^{+1} \to Q^+ \to 0$, we obtain $K_0(Q^+) \cong \mathbb{Z} \cong K_1(Q^+)$, where $K_0(Q^+)$ is generated by $[1_{Q^+}]$. All this also holds for the quotient Q^- of $E_1^0 \otimes E_n^{-1}$ by the ideal $\mathcal{K} \otimes \mathcal{K} \cong \mathcal{K}$.

Lemma 3.7 *Q^+ and Q^- are isomorphic. Moreover, there exists an automorphism of Q^+ which induces $\mathrm{id}_{\mathbb{Z}}$ on K_0 and $-\mathrm{id}_{\mathbb{Z}}$ on K_1.*

Proof: The first statement is an application of Theorem 3.5 to $A_1 = Q^+$, $A_2 = Q^-$, $I_1 = \mathcal{K} \otimes \mathcal{O}_n \triangleleft Q^+$, $J_1 = \mathcal{O}_1 \otimes \mathcal{K} \triangleleft Q^+$, $I_2 = \mathcal{K} \otimes \mathcal{O}_n \triangleleft Q^-$, $J_2 = \mathcal{O}_1 \otimes \mathcal{K} \triangleleft Q^-$. Namely, it is straightforward to check that it is possible to choose α_I, α_J, $\alpha_{I \oplus J}$, β, γ_I, γ_J, and $\gamma_{I \oplus J}$ with all the desired properties in Theorem 3.5.
The second statement follows in a similar way by applying Theorem 3.5 to $A_1 = A_2 = Q^+$, $I_1 = I_2 = \mathcal{K} \otimes \mathcal{O}_n \triangleleft Q^+$, $J_1 = J_2 = \mathcal{O}_1 \otimes \mathcal{K} \triangleleft Q^+$. □

Lemma 3.8 *For every $n \geq 2$, we have $E_1^0 \otimes E_n^{+1} \cong E_1^0 \otimes E_n^{-1}$.*

Proof: By the previous lemma, we can identify Q^+ and Q^- (we use the same notation as in the previous lemma) so that we can view $E_1^0 \otimes E_n^{+1}$ and $E_1^0 \otimes E_n^{-1}$ as extensions of Q^+:

$$0 \to \mathcal{K} \to E_1^0 \otimes E_n^{+1} \to Q^+ \to 0 \tag{2}$$

$$0 \to \mathcal{K} \to E_1^0 \otimes E_n^{-1} \to Q^+ \to 0. \tag{3}$$

Again by the previous lemma, we can choose the identification $Q^+ \cong Q^-$ in such a way that for a fixed choice of isomorphisms $K_1(Q^+) \cong \mathbb{Z}$, $K_0(\mathcal{K}) \cong \mathbb{Z}$, the index maps for both extensions (2) and (3) coincide. Now [BD96, Theorem 2] implies that (2) and (3) give the same class in $\mathrm{Ext}_s(Q^+)$. The reason is that $\mathrm{Ext}(K_0(Q^+), [1_{Q^+}])$ is the trivial group as $K_0(Q^+) \cong \mathbb{Z}$ and $[1_{Q^+}]$ is a generator of $K_0(Q^+) \cong \mathbb{Z}$. So the short exact sequence in [BD96, Theorem 2] tells us that two extensions of Q^+ by $\mathcal{K}$ give the same class in $\mathrm{Ext}_s(Q^+)$ if their index maps coincide. But this is the case for (2) and (3) by construction. Hence $E_1^0 \otimes E_n^{+1} \cong E_1^0 \otimes E_n^{-1}$ by [JT91, § 3.2]. □
For $m, n \geq 2$, the C*-algebra $E_m^{+1} \otimes E_n^{-1}$ contains the ideal $\mathcal{K} \otimes \mathcal{K} \cong \mathcal{K}$, and we denote the corresponding quotient by Q^{+-}. Obviously, the primitive ideals of Q^{+-} are given by $\mathcal{K} \otimes \mathcal{O}_n$, $\mathcal{O}_m \otimes \mathcal{K}$ and $\mathcal{K} \otimes \mathcal{O}_n \oplus \mathcal{O}_m \otimes \mathcal{K}$. From the six term exact sequence in K-theory for $0 \to \mathcal{K} \to E_m^{+1} \otimes E_n^{-1} \to Q^{+-} \to 0$, we obtain $K_0(Q^{+-}) \cong \mathbb{Z}/(m-1)(n-1)\mathbb{Z}$, with the class of the unit being a generator, and $K_1(Q^{+-}) \cong \{0\}$. All this also holds for the quotient Q^{-+} of $E_m^{-1} \otimes E_n^{+1}$ by the ideal $\mathcal{K} \otimes \mathcal{K} \cong \mathcal{K}$.

Lemma 3.9 *Q^{+-} and Q^{-+} are isomorphic.*

Proof: As Lemma 3.7, this is an application of Theorem 3.5 to $A_1 = Q^{+-}$, $A_2 = Q^{-+}$, $I_1 = \mathcal{K} \otimes \mathcal{O}_n \triangleleft Q^{+-}$, $J_1 = \mathcal{O}_m \otimes \mathcal{K} \triangleleft Q^{+-}$, $I_2 = \mathcal{K} \otimes \mathcal{O}_n \triangleleft Q^{-+}$,

$J_2 = \mathcal{O}_m \otimes \mathcal{K} \triangleleft Q^{-+}$. Namely, it is straightforward to check that it is possible to choose α_I, α_J, $\alpha_{I\oplus J}$, β, γ_I, γ_J, and $\gamma_{I\oplus J}$ with all the desired properties in Theorem 3.5. □

Lemma 3.10 *We have* $E_m^{+1} \otimes E_n^{-1} \cong E_m^{-1} \otimes E_n^{+1}$.

Proof: By the previous lemma, we can identify Q^{+-} and Q^{-+} (using the same notation as in the previous lemma) so that we can view $E_m^{+1} \otimes E_n^{-1}$ and $E_m^{-1} \otimes E_n^{+1}$ as extensions of Q^{+-}:

$$0 \to \mathcal{K} \to E_m^{+1} \otimes E_n^{-1} \to Q^{+-} \to 0 \tag{4}$$
$$0 \to \mathcal{K} \to E_m^{-1} \otimes E_n^{+1} \to Q^{+-} \to 0. \tag{5}$$

Since $\mathrm{Hom}(K_1(Q^{+-}), \mathbb{Z}) = \{0\}$, [BD96, Theorem 2] yields $\mathrm{Ext}((K_0(Q^{+-}), [1]), \mathbb{Z}) \cong Ext_s(Q^{+-})$. Hence (2) and (3) give the same class in $\mathrm{Ext}_s(Q^{+-})$. The reason is that the exact sequences in K_0 for (4) and (5) clearly give rise to the same class in $\mathrm{Ext}((K_0(Q^{+-}), [1]), \mathbb{Z})$. Hence $E_m^{+1} \otimes E_n^{-1} \cong E_m^{-1} \otimes E_n^{+1}$ by [JT91, § 3.2]. □

In an entirely analogous way, we get

Lemma 3.11 *For all* $m, n \geq 2$, *we have* $E_m^{+1} \otimes E_n^{+1} \cong E_m^{-1} \otimes E_n^{-1}$.

4 Classification of semigroup C*-algebras

We are now ready to address the general classification problem for C*-algebras of the form $C^*(A_\Gamma^+)$. We begin with notation:

Definition 4.1 *Let* Γ *be a graph with co-irreducible components* $\Gamma_i = (V_i, E_i)$. *We set*

$$t(\Gamma) = \#\{\Gamma_i\colon |V_i| = 1\}$$
$$o(\Gamma) = \#\{\Gamma_i\colon |V_i| = \infty\},$$

and for every $n \in \mathbb{Z}$

$$N_n(\Gamma) = \#\{\Gamma_i\colon 1 < |V_i| < \infty, \chi(\Gamma_i) = n\}.$$

Theorem 4.2 *Let* Γ *and* Λ *be two graphs. The semigroup C*-algebras* $C^*(A_\Gamma^+)$ *and* $C^*(A_\Lambda^+)$ *of the Artin monoids for* Γ *and* Λ *are stably isomorphic if and only if the following conditions hold:*

(i) $t(\Gamma) = t(\Lambda)$;

(ii) $N_{-n}(\Gamma) + N_n(\Gamma) = N_{-n}(\Lambda) + N_n(\Lambda)$ *for any* $n \in \mathbb{Z}$;

(iii) $\sum_{n\in\mathbb{Z}} N_n(\Gamma) = \infty$ *or* $\min(o(\Gamma), 1) = \min(o(\Lambda), 1)$.

They are isomorphic if and only if further *(iv) If $\sum_{n\in\mathbb{Z}} N_n(\Gamma) < \infty$, $o(\Gamma) = 0$ and $N_0(\Gamma) = 0$, then*

$$\sum_{n=1}^{\infty} N_{-n}(\Gamma) \equiv \sum_{n=1}^{\infty} N_{-n}(\Lambda) \mod 2$$

holds.

Remark 4.3 Note that when (ii) holds, all the conditions in (iii) are symmetric in Γ and Λ. Similarly, when (ii) and (iii) hold, so are the conditions in (iv).

For the proof of Theorem 4.2, we need some preparation. Given a graph Γ with co-irreducible components $\Gamma_i = (V_i, E_i)$, let Γ' be the graph we get from Γ by removing all the co-irreducible components Γ_i with $|V_i| = 1$ and the corresponding edges. We then have a canonical isomorphism $C^*(A_{\Gamma'}^+) \cong \bigotimes_{\{\Gamma_i\colon |V_i|>1\}} C^*(A_{\Gamma_i}^+)$.

Lemma 4.4 *There is a primitive ideal I' of $C^*(A_\Gamma^+)$ such that $\mathrm{Prim}\,(C^*(A_\Gamma^+)/I')$ does not continuously surject onto $\mathrm{Prim}\,(\mathcal{T})$ and which is minimal among all the primitive ideals having this property, and we have*

$$C^*(A_\Gamma^+)/I' \cong C^*(A_{\Gamma'}^+).$$

Proof: Let I be a primitive ideal of $C^*(A_\Gamma^+)$. As seen in Section 2, we know that I is generated by $\left\{\bigotimes_j J_{ij}\right\}_i$, where $J_{ij} = C^*(A_{\Gamma_j}^+)$ for $i \neq j$ and $J_{ii} = I_i$ for primitive ideals I_i of $C^*(A_{\Gamma_i}^+)$. It follows that $C^*(A_\Gamma^+)/I \cong \bigotimes_i C^*(A_{\Gamma_i}^+)/I_i$, and hence $\mathrm{Prim}\,(C^*(A_\Gamma^+)/I) \cong \prod_i \mathrm{Prim}\,(C^*(A_{\Gamma_i}^+)/I_i)$. We now claim that there exists a continuous surjection $\mathrm{Prim}\,(C^*(A_\Gamma^+)/I) \twoheadrightarrow \mathrm{Prim}\,\mathcal{T}$ if and only if there exists a co-irreducible component Γ_i of Γ with $|V_i| = 1$ and $I_i = (0)$. The direction "$\Leftarrow$" is obvious. For "$\Rightarrow$", assume that for every co-irreducible component Γ_i of Γ with $|V_i| = 1$, I_i is a maximal ideal of $C^*(A_{\Gamma_i}^+)$ such that $C^*(A_{\Gamma_i}^+)/I_i \cong \mathbb{C}$. Then $\mathrm{Prim}\,(C^*(A_\Gamma^+)/I) \cong \prod_k X_k$ where $X_k = \{x_k, y_k\}$ and the open subsets of X_k are given by $\emptyset$, $\{x_k\}$ and X_k. This means that $\overline{\{x_k\}} = X_k$ and $\overline{\{y_k\}} = \{y_k\}$. Furthermore, we know that $\mathrm{Prim}\,(\mathcal{T}) = \{\bullet\} \sqcup \mathbb{T}$, where $\overline{\{\bullet\}} = \mathrm{Prim}\,(\mathcal{T})$. Let $f : \prod_k X_k \to \mathrm{Prim}\,(\mathcal{T})$ be a continuous map. We want to show that f cannot be surjective. Let $y = (y_k)_k$ and $f(y) = z$. For arbitrary $x \in \prod_k X_k$, we always have $y \in \overline{\{x\}}$. As $f^{-1}(\overline{\{f(x)\}})$ is closed and contains x, it must also contain y. Hence $z = f(y)$ lies in $\overline{\{f(x)\}}$. This implies that $f(x) = z$ or $f(x) = \bullet$. But this holds for every x in $\prod_k X_k$. Hence the image of f contains at most 2 points, and thus f cannot be surjective. This shows our claim.
Therefore, a primitive ideal I' of $C^*(A_\Gamma^+)$ such that $\mathrm{Prim}\,(C^*(A_\Gamma^+)/I')$ does not continuously surject onto $\mathrm{Prim}\,(\mathcal{T})$ and which is minimal among all the primitive ideals with this property is generated by $\left\{\bigotimes_j J_{ij}\right\}_i$, where for a co-irreducible component Γ_i with $|V_i| = 1$, $J_{ii} = I_i$ is a maximal ideal of $C^*(A_{\Gamma_i}^+)$ with $C^*(A_{\Gamma_i}^+)/I_i \cong \mathbb{C}$, and for a co-irreducible component Γ_i with $|V_i| > 1$, $J_{ii} = (0)$. We conclude that $C^*(A_\Gamma^+)/I' \cong C^*(A_{\Gamma'}^+)$. □

Lemma 4.5 *Let Γ and Λ be two graphs.*

(1) If $C^*(A_\Gamma^+)$ and $C^*(A_\Lambda^+)$ are isomorphic, then $t(\Gamma) = t(\Lambda)$ and $C^*(A_{\Gamma'}^+) \cong C^*(A_{\Lambda'}^+)$.

(2) If $C^*(A_\Gamma^+) \otimes \mathcal{K}$ and $C^*(A_\Lambda^+) \otimes \mathcal{K}$ are isomorphic, then $t(\Gamma) = t(\Lambda)$ and $C^*(A_{\Gamma'}^+) \otimes \mathcal{K} \cong C^*(A_{\Lambda'}^+) \otimes \mathcal{K}$.

Proof: We first prove (1). Since an isomorphism $C^*(A_\Gamma^+) \cong C^*(A_\Lambda^+)$ sends the primitive ideal I to a primitive ideal of $C^*(A_\Lambda^+)$ with the analogous property, we conclude that every isomorphism $C^*(A_\Gamma^+) \cong C^*(A_\Lambda^+)$ induces an isomorphism $C^*(A_{\Gamma'}^+) \cong C^*(A_{\Lambda'}^+)$. To prove that $t(\Gamma) = t(\Lambda)$, we observe that the primitive ideals of $C^*(A_\Gamma^+)$ which are contained in I are in one-to-one correspondence with the subsets of $\{\Gamma_i\colon |V_i| = 1\}$. Again, as an isomorphism $C^*(A_\Gamma^+) \cong C^*(A_\Lambda^+)$ sends the primitive ideal I to a primitive ideal of $C^*(A_\Lambda^+)$ with the analogous property, we conclude that the power sets of $\{\Gamma_i\colon |V_i| = 1\}$ and $\{\Lambda_j\colon |W_j| = 1\}$ have the same cardinality. Hence also $\{\Gamma_i\colon |V_i| = 1\}$ and $\{\Lambda_j\colon |W_j| = 1\}$ must have the same cardinality (which is either finite or countably infinite). This proves (1).
(2) is proved in a similar way as (1) using the observation that every primitive ideal of $B \otimes \mathcal{K}$ is of the form $I \otimes \mathcal{K}$, where I is a primitive ideal of B. □

Lemma 4.6 *Let A_i, $i = 1, 2, \ldots$, be a countably infinite family of properly infinite unital C^*-algebras. Then $A = \bigotimes_{i=1}^\infty A_i$ is purely infinite.*

Proof: We have to show that every non-zero positive element a of A is properly infinite. By [KR02, Lemma 3.3], it suffices to find for every $\varepsilon > 0$ a properly infinite, positive element $b \in A$ with $\|a - b\| < \varepsilon$ and $b \precsim a$. Since $A = \bigotimes_{i=1}^\infty A_i$, there exists a (sufficiently large) natural number n and a positive element $x \in \bigotimes_{i=1}^n A_i$ with $\|a - x \otimes 1\| < \frac{\varepsilon}{2}$. By [KR02, Lemma 2.2], we have that $b := (x - \frac{\varepsilon}{2})_+ \otimes 1 = (x \otimes 1 - \frac{\varepsilon}{2})_+$ satisfies $b \precsim a$. Also, we have $\|b - a\| \leq \|b - x \otimes 1\| + \|x \otimes 1 - a\| < \varepsilon$. So it suffices to show that b is properly infinite. By construction, b is of the form $c \otimes 1$ for some positive element $c \in \bigotimes_{i=1}^n A_i$. Since the unit $1 \in A_{n+1}$ is properly infinite, we can find isometries s and t in A_{n+1} with $ss^* \perp tt^*$. So $b = c \otimes 1 = (c^{1/2} \otimes s)^*(c^{1/2} \otimes s) \approx (c^{1/2} \otimes s)(c^{1/2} \otimes s)^* = c \otimes ss^*$. Similarly, $b \approx c \otimes tt^*$. But since $(c \otimes ss^*)(c \otimes tt^*) = 0$, we conclude that $b \oplus b \approx (c \otimes ss^*) \oplus (c \otimes tt^*) \approx c \otimes (ss^* + tt^*) \leq c \otimes 1 = b$. □

Lemma 4.7 *Let Γ be a graph with (countably) infinitely many co-irreducible components $\Gamma_i = (V_i, E_i)$, $i = 1, 2, \ldots$. Assume that $1 < |V_i| < \infty$ for all i. Then $C^*(A_\Gamma^+)$ is strongly purely infinite, i.e., $C^*(A_\Gamma^+) \cong C^*(A_\Gamma^+) \otimes \mathcal{O}_\infty$.*

Proof: By [CL07, Theorem 8.3], we know that $C^*(A_\Gamma^+)$ has the ideal property (the definition can be found in [PR07, Remark 2.1]). Moreover, we know that $C^*(A_\Gamma^+) \cong \bigotimes_{i=1}^\infty C^*(A_{\Gamma_i}^+)$, and each of the $C^*(A_{\Gamma_i}^+)$ is a properly infinite unital C*-algebra. Hence by the previous lemma, we know that $C^*(A_\Gamma^+)$ is purely infinite. Therefore, [PR07, Proposition 2.14] tells us that $C^*(A_\Gamma^+)$ is strongly purely infinite. And finally, if $C^*(A_\Gamma^+)$ is strongly purely infinite, then [KR02, Theorem 9.1] implies that $C^*(A_\Gamma^+) \cong C^*(A_\Gamma^+) \otimes \mathcal{O}_\infty$ because $C^*(A_\Gamma^+)$ is nuclear and unital. □

Finally, we are ready for the proof of Theorem 4.2. *Proof:*[Proof of Theorem 4.2] Let us first of all show that if $C^*(A_\Gamma^+) \otimes \mathcal{K} \cong C^*(A_\Lambda^+) \otimes \mathcal{K}$ holds, then conditions (i), (ii) and (iii) must be satisfied. By Lemma 4.5 condition (i) holds and that $C^*(A_{\Gamma'}^+) \otimes \mathcal{K} \cong C^*(A_{\Lambda'}^+) \otimes \mathcal{K}$. Hence we may assume that all the co-irreducible components of Γ and Λ have more than one vertex.

To prove (ii), we observe that the minimal non-zero primitive ideals of $C^*(A_\Gamma^+)$ are of the form $I_i = \otimes_j J_{ij}$, where $J_{ij} = C^*(A_{\Gamma_j}^+)$ if $j \neq i$, and $J_{ii} = \mathcal{K} \triangleleft C^*(A_{\Gamma_i}^+)$ (Γ_i consists of only finitely many vertices). For the corresponding quotient, we get $C^*(A_\Gamma^+)/I_i \cong \bigotimes_j Q_{ij}$, where $Q_{ij} = C^*(A_{\Gamma_j}^+)$ if $j \neq i$, and $Q_{ii} = C^*(A_{\Gamma_i}^+)/\mathcal{K}$. Since $K_0(C^*(A_{\Gamma_i}^+)/\mathcal{K}) \cong \mathbb{Z}/|\chi(\Gamma_i)|\mathbb{Z}$, it follows that $K_0(C^*(A_\Gamma^+)/I_i) \cong \mathbb{Z}/|\chi(\Gamma_i)|\mathbb{Z}$. Hence, we have shown that $N_0(\Gamma)$ is the number of minimal non-zero primitive ideals $I \otimes \mathcal{K}$ of $C^*(A_\Gamma^+) \otimes \mathcal{K}$ with the property that $K_0(C^*(A_\Gamma^+)/I) \cong \mathbb{Z}$, and that for every $n = 1, 2, \dots,$ $N_{-n}(\Gamma) + N_n(\Gamma)$ is the number of minimal non-zero primitive ideals $I \otimes \mathcal{K}$ of $C^*(A_\Gamma^+) \otimes \mathcal{K}$ with the property that $K_0(C^*(A_\Gamma^+)/I) \cong \mathbb{Z}/n\mathbb{Z}$. Since these descriptions are invariant under stable isomorphisms of C*-algebras, we conclude that (ii) must hold.

Let us now prove (iii) under the assumption of stable isomorphism. If $\sum N_n(\Gamma) = \infty$ we are done, so suppose the contrary and note that in this case, $C^*(A_\Gamma^+)$ is strongly purely infinite if and only if $o(\Gamma) > 0$. The direction "$\Rightarrow$" is clear, since $o(\Gamma) > 0$ implies that $C^*(A_\Gamma^+)$ has $\mathcal{O}_\infty$ as a tensor factor. To prove "$\Leftarrow$", we observe that if $o(\Gamma) = 0$, then $C^*(A_\Gamma^+)$ contains the algebra of compact operators as an ideal, hence cannot be strongly purely infinite. As a consequence, $C^*(A_\Gamma^+) \otimes \mathcal{K} \cong C^*(A_\Lambda^+) \otimes \mathcal{K}$ implies that either both $o(\Gamma) > 0$ and $o(\Lambda) > 0$, or $o(\Gamma) = o(\Lambda) = 0$, as desired.

Finally, we assume that $C^*(A_\Gamma^+) \cong C^*(A_\Lambda^+)$ and that $\sum N_n(\Gamma) < \infty$, that $N_0(\Gamma) = 0$ and that $o(\Gamma) = 0$. The algebra $\mathcal{K}$ of compact operators sits inside $C^*(A_\Gamma^+)$ as the (unique) minimal non-zero ideal. The inclusion $\mathcal{K} \hookrightarrow C^*(A_\Gamma^+)$ sends in K-theory the K_0-class of a minimal projection to $(\prod_i \chi(\Gamma_i)) \cdot [1]$, where $(\prod_i \chi(\Gamma_i))$ is the product over all co-irreducible components of Γ (there are only finitely many by assumption) of the Euler characteristics. As $N_0(\Gamma) = 0$, $(\prod_i \chi(\Gamma_i))$ is a non-zero number, and it is positive if and only if $\sum_{n=1}^\infty N_{-n}(\Gamma) \equiv 0 \mod 2$. Since $C^*(A_\Gamma^+) \cong C^*(A_\Lambda^+)$, we must have $\sum_{n=1}^\infty N_{-n}(\Gamma) \equiv \sum_{n=1}^\infty N_{-n}(\Lambda) \mod 2$. Therefore, all in all, condition (iv) follows when the C*-algebras are isomorphic.

In the opposite direction, we know from Sections 2 and 3 that

$$C^*(A_\Gamma^+) \cong \mathcal{T}^{\otimes t(\Gamma)} \otimes \mathcal{O}_\infty^{\otimes o(\Gamma)} \otimes \bigotimes_{n=0}^{\infty} \bigotimes_{\{i:\ |\chi(\Gamma_i)|=n\}} E_{1+n}^{\mathrm{sgn}(\chi(\Gamma_i))}$$

and

$$C^*(A_\Lambda^+) \cong \mathcal{T}^{\otimes t(\Lambda)} \otimes \mathcal{O}_\infty^{\otimes o(\Lambda)} \otimes \bigotimes_{n=0}^{\infty} \bigotimes_{\{i:\ |\chi(\Lambda_i)|=n\}} E_{1+n}^{\mathrm{sgn}(\chi(\Lambda_i))}$$

We note from the outset that whenever $o(\Gamma) > 0$ or $N_0(\Gamma) > 0$ then by repeated application of either Lemma 3.6 or Lemma 3.8 we may simplify these expressions to

$$C^*(A_\Gamma^+) \cong \mathcal{T}^{\otimes t(\Gamma)} \otimes \mathcal{O}_\infty^{\otimes o(\Gamma)} \otimes (E_1^0)^{\otimes N_0(\Gamma)} \otimes \bigotimes_{n=1}^{\infty} (E_{1+n}^{+1})^{\otimes (N_{-n}(\Gamma)+N_n(\Gamma))}. \quad (6)$$

Assume that (i), (ii) and (iii) hold. We begin by noting that in the case $\sum N_n(\Gamma) < \infty$ if either $o(\Gamma) > 0$ or $N_0(\Gamma) > 0$, we also have either $o(\Lambda) > 0$ or $N_0(\Lambda) > 0$, and we get $C^*(A_\Gamma^+) \cong C^*(A_\Lambda^+)$ by reducing to the form given in (6) and applying (i) and (ii).

When $\sum N_n(\Gamma) = \infty$ then we have by (ii) and Lemma 4.7 that both $C^*(A_\Gamma^+)$ and $C^*(A_\Lambda^+)$ are strongly purely infinite, and hence we have

$$\begin{aligned} C^*(A_\Gamma^+) &\cong \mathcal{T}^{\otimes t(\Gamma)} \otimes \mathcal{O}_\infty \otimes (E_1^0)^{\otimes N_0(\Gamma)} \otimes \bigotimes_{n=1}^{\infty} (E_{1+n}^{+1})^{\otimes (N_{-n}(\Gamma)+N_n(\Gamma))} \\ &= \mathcal{T}^{\otimes t(\Lambda)} \otimes \mathcal{O}_\infty \otimes (E_1^0)^{\otimes N_0(\Lambda)} \otimes \bigotimes_{n=1}^{\infty} (E_{1+n}^{+1})^{\otimes (N_{-n}(\Lambda)+N_n(\Lambda))} \\ &\cong C^*(A_\Lambda^+), \end{aligned}$$

since Lemma 3.6 may be applied as above.

It remains to treat the case that $o(\Gamma) = N_0(\Gamma) = 0$ and $\sum N_n(\Gamma) < \infty$. Again, by (ii) and (iii), we must have $o(\Lambda) = N_0(\Lambda) = 0$ and $\sum N_n(\Lambda) < \infty$ as well, and we get

$$\begin{aligned} C^*(A_\Gamma^+) \otimes \mathcal{K} &\cong \mathcal{T}^{\otimes t(\Gamma)} \otimes \bigotimes_{n=1}^{\infty} (E_{1+n}^{+1})^{\otimes (N_{-n}(\Gamma)+N_n(\Gamma))} \otimes \mathcal{K} \\ &= \mathcal{T}^{\otimes t(\Lambda)} \otimes \bigotimes_{n=1}^{\infty} (E_{1+n}^{+1})^{\otimes (N_{-n}(\Lambda)+N_n(\Lambda))} \otimes \mathcal{K} \\ &\cong C^*(A_\Lambda^+) \otimes \mathcal{K} \end{aligned}$$

this time appealing to the second half of Lemma 3.3.

Assuming further (iv), we now aim for exact isomorphism, noting that we have already established it when $o(\Gamma) > 0$, $N_0(\Gamma) > 0$ or $\sum N_n(\Gamma) = \infty$. We hence assume that $o(\Gamma) = N_0(\Gamma) = 0$ and note that also $o(\Lambda) = N_0(\Lambda) = 0$

Consider first the case where both $\sum_{n=1}^{\infty} N_{-n}(\Gamma)$ and $\sum_{n=1}^{\infty} N_{-n}(\Lambda)$ are even. We have

$$\begin{aligned} C^*(A_\Gamma^+) &\cong \mathcal{T}^{\otimes t(\Gamma)} \otimes \bigotimes_{n=1}^{\infty} (E_{1+n}^{+1})^{\otimes N_n(\Gamma)} \otimes \bigotimes_{n=1}^{\infty} (E_{1+n}^{-1})^{\otimes N_{-n}(\Gamma)} \\ &\cong \mathcal{T}^{\otimes t(\Gamma)} \otimes \bigotimes_{n=1}^{\infty} (E_{1+n}^{+1})^{\otimes N_n(\Gamma)} \otimes \bigotimes_{n=1}^{\infty} (E_{1+n}^{+1})^{\otimes N_{-n}(\Gamma)} \\ &\cong \mathcal{T}^{\otimes t(\Lambda)} \otimes \bigotimes_{n=1}^{\infty} (E_{1+n}^{+1})^{\otimes (N_{-n}(\Gamma)+N_n(\Gamma))} \\ &= \mathcal{T}^{\otimes t(\Lambda)} \otimes \bigotimes_{n=1}^{\infty} (E_{1+n}^{+1})^{\otimes (N_{-n}(\Lambda)+N_n(\Lambda))} \cong C^*(A_\Lambda^+) \end{aligned}$$

by Lemma 3.11. Now assume both $\sum_{n=1}^{\infty} N_{-n}(\Gamma)$ and $\sum_{n=1}^{\infty} N_{-n}(\Lambda)$ are odd. If there exists $\chi < 0$ such that there are co-irreducible components Γ_k and Λ_l with $\chi(\Gamma_k) = \chi = \chi(\Lambda_l)$, then we deduce from the previous case that

$$\begin{aligned} C^*(A_\Gamma^+) &\cong \left(\bigotimes_{\Gamma_i \neq \Gamma_k} C^*(A_{\Gamma_i}^+) \right) \otimes C^*(A_{\Gamma_k}^+) \\ &\cong \left(\bigotimes_{\Gamma_i \neq \Gamma_k} C^*(A_{\Gamma_i}^+) \right) \otimes E_{1+|\chi|}^{-1} \\ &\cong \left(\bigotimes_{\Lambda_j \neq \Lambda_l} C^*(A_{\Lambda_j}^+) \right) \otimes C^*(A_{\Lambda_l}^+) \cong C^*(A_\Lambda^+). \end{aligned}$$

If there exists no such χ, then by (ii) there must be $\chi < 0$, $\psi < 0$ and co-irreducible components Γ_{k_-}, Γ_{k_+}, Λ_{l_-}, Λ_{l_+} with $\chi(\Gamma_{k_-}) = \chi$, $\chi(\Lambda_{l_+}) = -\chi$, $\chi(\Gamma_{k_+}) = -\psi$ and $\chi(\Lambda_{l_-}) = \psi$. Hence

$$\begin{aligned} C^*(A_\Gamma^+) &\cong \left(\bigotimes_{\Gamma_i \neq \Gamma_{k_+}, \Gamma_{k_-}} C^*(A_{\Gamma_i}^+) \right) \otimes C^*(A_{\Gamma_{k_+}}^+) \otimes C^*(A_{\Gamma_{k_-}}^+) \\ &\cong \left(\bigotimes_{\Gamma_i \neq \Gamma_{k_+}, \Gamma_{k_-}} E_{1+|\chi(\Gamma_i)|}^{\mathrm{sgn}(\chi(\Gamma_i))} \right) \otimes E_{1+|\psi|}^{+1} \otimes E_{1+|\chi|}^{-1} \\ &\cong \left(\bigotimes_{\Gamma_i \neq \Gamma_{k_+}, \Gamma_{k_-}} E_{1+|\chi(\Gamma_i)|}^{\mathrm{sgn}(\chi(\Gamma_i))} \right) \otimes E_{1+|\psi|}^{-1} \otimes E_{1+|\chi|}^{+1} \\ &\cong \left(\bigotimes_{\Gamma_i \neq \Gamma_{k_+}, \Gamma_{k_-}} C^*(A_{\Gamma_i}^+) \right) \otimes C^*(A_{\Lambda_{l_-}}^+) \otimes C^*(A_{\Lambda_{l_+}}^+) \cong C^*(A_\Lambda^+). \end{aligned}$$

In the third step, we used Lemma 3.10, and in the fourth step, we used our argument in the previous case. $\square$

5 The isomorphism problem from the perspective of classification of non-simple C*-algebras

We give an interpretation of Theorem 4.2 from the point of view of classifying non-simple C*-algebras.
We let $\mathcal{O}(\mathrm{Prim}(A))$ denote the set of open subsets in $\mathrm{Prim}(A)$, and $\mathbb{I}(A)$ the lattice of ideals.
A lattice map $\psi_A\colon \mathcal{O}(\mathrm{Prim}(A)) \to \mathbb{I}(A)$ given by $\psi_A(U) = \bigcap_{\rho \notin U} \rho$ is then a lattice isomorphism which preserves arbitrary suprema and finite infima. We denote $\psi_A(U)$ by $A[U]$. For every C*-algebra A, we denote the pair

$$\left(\mathrm{Prim}(A), \{\mathrm{K}^+_{\mathrm{six}}(A/A[U]; A[V]/A[U])\}_{\substack{V,U \in \mathbb{O}(\mathrm{Prim}(A)) \\ U \subseteq V}}\right)$$

by $\mathfrak{F}(A)$, where $\mathrm{K}^+_{\mathrm{six}}(B, J)$ denotes the standard six-term exact sequence associated to an ideal J of a C*-algebra B, considering each K_0-group as an ordered group.
An isomorphism from $\mathfrak{F}(A)$ to $\mathfrak{F}(B)$ thus consists of a homeomorphism

$$\phi\colon \mathrm{Prim}(A) \to \mathrm{Prim}(B)$$

and isomorphisms

$$\alpha_{U,V}\colon K_*(A[V]/A[U]) \to K_*(B[\phi(V)]/B[\phi(U)])$$

for each $U, V \in \mathbb{O}(\mathrm{Prim}(A))$ with $U \subseteq V$, such that $(\alpha_{U,V}, \alpha_{X,U}, \alpha_{X,V})$ is an isomorphism from $\mathrm{K}^+_{\mathrm{six}}(A/A[U]; A[V]/A[U])$ to

$$\mathrm{K}^+_{\mathrm{six}}(B/B[\phi(U)]; B[\phi(V)]/B[\phi(U)])$$

in the sense that it makes all squares commute and is an order isomorphism on all even parts of the K-theory.
If A and B are unital, we write $(\mathfrak{F}(A), [1_A]) \cong (\mathfrak{F}(B), [1_B])$ if $\mathfrak{F}(A) \cong \mathfrak{F}(B)$ in such a way that the isomorphism $\alpha_{X,\emptyset}$ sends $[1_A]$ in $K_0(A)$ to $[1_B]$ in $K_0(B)$.
Note that if $\phi\colon \mathrm{Prim}(A) \to \mathrm{Prim}(B)$ is a homeomorphism, there exists a lattice isomorphism from $\mathbb{I}(A)$ to $\mathbb{I}(B)$ given by $I \mapsto \psi_B(\phi(\psi_A^{-1}(I)))$. Hence, if A and B are separable and $\phi\colon \mathrm{Prim}(A) \to \mathrm{Prim}(B)$ is a homeomorphism, then for all $U \in \mathbb{O}(\mathrm{Prim}(A))$, we have that $A[U]$ is a primitive ideal of A if and only if $B[\phi(U)]$ is a primitive ideal of B (because primitive ideals are precisely given by prime ideals for separable C*-algebras).
The following easy observation is left to the reader.

Lemma 5.1 *Let A and B be separable C*-algebras. Let $U \in \mathbb{O}(\mathrm{Prim}(A))$.*

(1) If $\mathfrak{F}(A) \cong \mathfrak{F}(B)$ *via a homeomorphism* $\phi\colon \mathrm{Prim}(A) \to \mathrm{Prim}(B)$*, then*

$$\mathfrak{F}(A/A[U]) \cong \mathfrak{F}(B/B[\phi(U)]).$$

(2) If A *and* B *are unital* C^**-algebras and* $(\mathfrak{F}(A),[1_A]) \cong (\mathfrak{F}(B),[1_B])$ *via a homeomorphism* $\phi\colon \mathrm{Prim}(A) \to \mathrm{Prim}(B)$*, then*

$$\left(\mathfrak{F}(A/A[U]),[1_{A/A[U]}]\right) \cong \left(\mathfrak{F}(B/B[\phi(U)]),[1_{B/B[\phi(U)]}]\right).$$

Theorem 5.2 *Let* Γ *and* Λ *be two (countable) graphs, and let* $\Gamma_i = (V_i, E_i)$ *be the co-irreducible components of* Γ*,* $\Lambda_j = (W_j, F_j)$ *the co-irreducible components of* Λ*. Then* $C^*(A_\Gamma^+) \cong C^*(A_\Lambda^+)$ *if and only if* $(\mathfrak{F}(C^*(A_\Gamma^+)),[1_{C^*(A_\Gamma^+)}]) \cong (\mathfrak{F}(C^*(A_\Lambda^+)),[1_{C^*(A_\Lambda^+)}])$.

Proof: The direction "$\Rightarrow$" is obvious. To prove "$\Leftarrow$", we show that

$$(\mathfrak{F}(C^*(A_\Gamma^+)),[1_{C^*(A_\Gamma^+)}]) \cong (\mathfrak{F}(C^*(A_\Lambda^+)),[1_{C^*(A_\Lambda^+)}])$$

implies (i), (ii), (iii) and (iv) from Theorem 4.2, using the notations from Lemma 4.5. The first step is to prove that $(\mathfrak{F}(C^*(A_\Gamma^+)),[1_{C^*(A_\Gamma^+)}]) \cong (\mathfrak{F}(C^*(A_\Lambda^+)),[1_{C^*(A_\Lambda^+)}])$ implies $t(\Gamma) = t(\Lambda)$ and $(\mathfrak{F}(C^*(A_{\Gamma'}^+)),[1_{C^*(A_{\Gamma'}^+)}]) \cong (\mathfrak{F}(C^*(A_{\Lambda'}^+)),[1_{C^*(A_{\Lambda'}^+)}])$. $t(\Gamma) = t(\Lambda)$ follows by Lemma 4.4, because we only use the primitive ideal space and the lattice structure of the set of ideals in this proof. To see that $(\mathfrak{F}(C^*(A_{\Gamma'}^+)),[1_{C^*(A_{\Gamma'}^+)}]) \cong (\mathfrak{F}(C^*(A_{\Lambda'}^+)),[1_{C^*(A_{\Lambda'}^+)}])$, let I' be a primitive ideal of $C^*(A_\Gamma^+)$ stipulated in Lemma 4.4, and let U be an open set of $\mathrm{Prim}\,(C^*(A_\Gamma^+))$ such that $C^*(A_\Gamma^+)[U] = I'$. Then $C^*(A_\Lambda^+)[\phi(U)]$ is an ideal with the analogous property. In the proof of Lemma 4.5, we have seen that $C^*(A_\Gamma^+)/C^*(A_\Gamma^+)[U] \cong C^*(A_{\Gamma'}^+)$. Similarly, we have $C^*(A_\Lambda^+)/C^*(A_\Lambda^+)[U] \cong C^*(A_{\Lambda'}^+)$. Therefore, (2) from the previous lemma tells us that $(\mathfrak{F}(C^*(A_{\Gamma'}^+)),[1_{C^*(A_{\Gamma'}^+)}]) \cong (\mathfrak{F}(C^*(A_{\Lambda'}^+)),[1_{C^*(A_{\Lambda'}^+)}])$, as desired.

In particular, this implies (i), and we may assume as in the proof of Theorem 4.2 that all the co-irreducible components of Γ and Λ have more than one vertex. Then (ii) follows in exactly the same way as in the proof of Theorem 4.2 because we only use primitive ideal spaces, lattice structures of sets of ideals and K_0 in this proof. All this can be extracted from the invariant $\mathfrak{F}$. Let us prove (iii). As we have seen in the proof of Theorem 4.2, $o(\Gamma) = 0$ implies that $\mathcal{K}$ is an ideal of $C^*(A_\Gamma^+)$, whereas $o(\Gamma) > 0$ implies that $C^*(A_\Gamma^+)$ (and hence also every non-zero ideal) is strongly purely infinite. These two cases can be distinguished by the order on K_0. Therefore, we see as in the proof of Theorem 4.2 that if $(\mathfrak{F}(C^*(A_\Gamma^+)),[1_{C^*(A_\Gamma^+)}]) \cong (\mathfrak{F}(C^*(A_\Lambda^+)),[1_{C^*(A_\Lambda^+)}])$, then either $o(\Gamma) > 0$ and $o(\Lambda) > 0$ or $o(\Gamma) = 0$ and $o(\Lambda) = 0$. The proof of (iv) then follows the proof of Theorem 4.2, where we only use lattice structures of sets of ideals, K_0 and the K_0-classes of the units. All this can be extracted from the invariant $\mathfrak{F}$ together with the position of the K_0-class of the unit. $\square$

6 Graph algebras and the semiprojectivity question

Apart from semigroup C*-algebras we discussed above, there is another - more traditional - way of constructing a C*-algebra out of a directed graph, possibly allowing for loops. Now we would like to discuss the overlap of these two constructions. In other words, we are interested in the question: Which semigroup C*-algebras for right-angled Artin monoids are isomorphic to graph algebras? We can provide a complete answer to this question.

6.1 Extensions of C*-algebras

We first establish some facts about absorbing extensions and the C*-algebras associated to these extensions. To each injective Busby map $\tau\colon A \to \mathcal{Q}(B)$, where $\mathcal{Q}(B) = \mathcal{M}(B)/B$ with $\mathcal{M}(B)$ the multiplier algebra of B, associate as usual the extension

$$e: \qquad \begin{tikzcd} 0 \arrow[r] & B \arrow[r, hook] \arrow[d, equal] & E \arrow[r, "\psi"] \arrow[d] & A \arrow[r] \arrow[d, "\tau"] & 0 \\ 0 \arrow[r] & B \arrow[r, hook] & \mathcal{M}(B) \arrow[r, "\pi"] & \mathcal{Q}(B) \arrow[r] & 0 \end{tikzcd}$$

with $E = \pi^{-1}(\tau(A))$ and $\psi(x) = \tau^{-1}(\pi(x))$. Note that ψ is a homomorphism since τ is injective.

We call τ (and e) *unital* if A is unital and τ is a unital homomorphism, or, equivalently, if E is a unital C*-algebra. If $\tau = \pi \circ \alpha$ for some homomorphism $\alpha\colon A \to \mathcal{M}(B)$, then τ is called a *trivial extension.* If A is unital and $\tau = \pi \circ \alpha$ for some unital homomorphism $\alpha\colon A \to \mathcal{M}(B)$, then τ is called *strongly unital.* Not all unital trivial extensions are strongly unital.

Assume that B is stable. The sum $\tau \oplus \tau'$ of two extensions $\tau, \tau'\colon A \to \mathcal{Q}(B)$ is defined as follows. Since B is stable, there exist isometries $s_1, s_2 \in \mathcal{M}(B)$ with $1_{\mathcal{M}(B)} = s_1 s_1^* + s_2 s_2^*$. Set

$$(\tau \oplus \tau')(a) = \pi(s_1)\tau(a)\pi(s_1^*) + \pi(s_2)\tau'(a)\pi(s_2^*)$$

for all $a \in A$.

Two extensions $\tau, \tau'\colon A \to \mathcal{Q}(B)$ are said to be *unitarily equivalent*, denoted by $\tau \sim_u \tau'$, if there exists a unitary $u \in \mathcal{M}(B)$ such that $\pi(u)\tau(a)\pi(u)^* = \tau'(a)$ for all $a \in A$. Then two extensions $\tau_1, \tau_2\colon A \to \mathcal{Q}(B)$ define the same element in $\mathrm{Ext}(A, B)$ if there exists a unitary $u \in \mathcal{M}(B)$ and there exist trivial extensions $\tau_1', \tau_2'\colon A \to \mathcal{Q}(B)$ such that $\tau_1 \oplus \tau_1' \sim_u \tau_2 \oplus \tau_2'$. If τ_1 and τ_2 are unital extensions, then τ_1' and τ_2' can be chosen to be unital extensions (see [Ror97, Section 5]).

For a C*-algebra C, we let $\widetilde{C}$ be the unitization of C (adding a new unit if C is a unital C*-algebra) and let $\iota_C\colon C \to \widetilde{C}$ be the embedding of C into $\widetilde{C}$ as an ideal.

Recall that an ideal I of a C*-algebra A is an *essential ideal* if every nonzero ideal of A has a nontrivial intersection with I. An extension $0 \to I \xrightarrow{\iota} A \to$

$B \to 0$ is *essential* if $\iota(I)$ is an essential ideal of A. It is a well-known fact that an extension $0 \to I \to A \to B \to 0$ is an essential extension if and only if the Busby invariant of the extension is injective. We now prove in the following proposition that every essential extension $0 \to B \to E \to A \to 0$ with A a non-unital, separable, nuclear C*-algebra and B a C*-algebra that is isomorphic to either $\mathcal{K}$ or a nuclear, purely infinite simple C*-algebra is absorbing.

Before proving the proposition, we show that any absorbing extension must be an essential extension. Hence, the assumption that the extension is essential is necessary. Note that if τ or τ' is injective, then the sum $\tau \oplus \tau'$ is injective. Since B is stable, there exists a unital embedding from $\mathcal{O}_2$ to $\mathcal{M}(B)$ which induces a unital embedding from $\mathcal{O}_2$ to $\mathcal{Q}(B)$. Nuclearity of A gives us an embedding of A into $\mathcal{O}_2$, thus the composition gives a trivial essential extension $\tau_0 \colon A \to \mathcal{Q}(B)$. Therefore, an absorbing extension τ is an essential extension since τ is unitarily equivalent to $\tau \oplus \tau_0$.

Proposition 6.1 *Let A be a non-unital, separable, nuclear C*-algebra and let B be a separable C*-algebra that is isomorphic to either $\mathcal{K}$ or a nuclear, purely infinite simple C*-algebra. If $\tau \colon A \to \mathcal{Q}(B)$ is an essential extension, then for every trivial extension $\tau_0 \colon A \to \mathcal{Q}(B)$ we have that $\tau \sim_u \tau \oplus \tau_0$. Consequently, if $e_i : 0 \to B \to E_i \to A \to 0$ is an essential extension for $i = 1, 2$ and $[\tau_{e_1}] = [\tau_{e_2}]$ in $\mathrm{Ext}(A, B)$, then $E_1 \cong E_2$.*

Proof: Let $\alpha_0 \colon A \to \mathcal{M}(B)$ be a homomorphism with $\tau_0 = \pi \circ \alpha_0$. Extend τ and α_0 to the unitization of $\widetilde{A}$, and denote these extensions by $\widetilde{\tau} \colon \widetilde{A} \to \mathcal{Q}(B)$ and $\widetilde{\alpha}_0 \colon \widetilde{A} \to \mathcal{M}(B)$ respectively.

We claim that $\widetilde{\tau}$ is injective. Let $y \in \ker(\widetilde{\tau})$. Then $\tau(yx) = \widetilde{\tau}(y)\widetilde{\tau}(\iota_A(x)) = 0$ and $\tau(xy) = \widetilde{\tau}(\iota_A(x))\widetilde{\tau}(y) = 0$ for all $x \in A$. Since τ is injective, we have that $yx = xy = 0$ for all $x \in A$. Since A is non-unital, A is an essential ideal of $\widetilde{A}$. Hence, $y = 0$. Thus, proving our claim.

Set $E = \pi^{-1}(\widetilde{\tau}(\widetilde{A})) \subseteq \mathcal{M}(B)$. Since $\widetilde{\tau}$ is injective, we may define a surjective homomorphism $\psi \colon E \to \widetilde{A}$ by $\psi(x) = \widetilde{\tau}^{-1}(\pi(x))$. Define $\eta \colon E \to \mathcal{M}(B)$ by $\eta(x) = \widetilde{\alpha}_0 \circ \psi(x)$. Then η is a unital homomorphism such that $\eta(E \cap B) = \{0\}$. Let s_1 and s_2 be isometries such that $1_{\mathcal{M}(B)} = s_1 s_1^* + s_2 s_2^*$. By [Arv77, Corollary 2] and [Kir, Proposition 7], there exists a unitary $u \in \mathcal{M}(B)$ such that $u(s_1 x s_1^* + s_2 \eta(x) s_2^*)u^* - x \in B$ for all $x \in E$.

We claim u implements a unitary equivalence between τ and $\tau \oplus \tau_0$. Let $a \in A$. Choose $x \in E$ such that $\pi(x) = (\widetilde{\tau} \circ \iota_A)(a)$. Note that

$$\pi \circ \eta(x) = \pi \circ \widetilde{\alpha}_0 \circ \psi(x) = (\pi \circ \widetilde{\alpha}_0)(\widetilde{\tau}^{-1}(\pi(x))) = (\pi \circ \widetilde{\alpha}_0 \circ \iota_A)(a).$$

Then

$$\begin{aligned}
&\pi(u)\,(\tau(a) \oplus \tau_0(a))\,\pi(u)^* \\
&\quad = \pi(u)\,(\pi(s_1)\widetilde{\tau}(\iota_A(a))\pi(s_1^*) + \pi(s_2)(\pi \circ \widetilde{\alpha}_0 \circ \iota_A)(a)\pi(s_2)^*)\,\pi(u)^* \\
&\quad = \pi(u(s_1 x s_1^* + s_2 \eta(x) s_2^*)u^*) \\
&\quad = \pi(x) \\
&\quad = (\widetilde{\tau} \circ \iota_A)(a) \\
&\quad = \tau(a).
\end{aligned}$$

Hence, $\tau \oplus \tau_0 \sim_u \tau$, proving the first part of the proposition.
Suppose $e_i : 0 \to B \to E_i \to A \to 0$ is an essential extension for $i = 1, 2$ and $[\tau_{e_1}] = [\tau_{e_2}]$ in $\mathrm{Ext}(A, B)$. By the discussion before the proposition, there exist trivial extensions $\tau_1', \tau_2' \colon A \to \mathcal{Q}(B)$ such that $\tau_{e_1} \oplus \tau_1' \sim_u \tau_{e_2} \oplus \tau_2'$. By the first part of the proposition, we have that $\tau_{e_1} \sim_u \tau_{e_1} \oplus \tau_1'$ and $\tau_{e_2} \oplus \tau_2' \sim_u \tau_{e_2}'$. Therefore, $\tau_{e_1} \sim_u \tau_{e_2}$. By [JT91, § 3.2], $E_1 \cong E_2$. □

6.2 Corners of graph algebras

We also need some results involving corners of graph algebras. The general case will be worked out in [AGR]. For the convenience of the reader, we will prove the case that will suit our purposes (see Proposition 6.2).
Recall that if $E = (E^0, E^1, r, s)$ is a graph, the C*-algebra $C^*(E)$ associated to E is the universal C*-algebra generated by $\{p_v : v \in E^0\} \sqcup \{s_e : e \in E^1\}$ subject to the relations

(i) $p_v p_w = \delta_{v,w} p_v$ for all $v, w \in E^0$;

(ii) $s_e^* s_f = \delta_{e,f} p_{r(e)}$ for all $e, f \in E^1$;

(iii) $s_e s_e^* \leq p_{s(e)}$ for all $e \in E^1$; and

(iv) $p_v = \sum_{e \in s^{-1}(v)} s_e s_e^*$ for all $v \in E^0$ with $0 < |s^{-1}(v)| < \infty$.

A *loop* in E is a path $\alpha = e_1 \cdots e_n$ with $s(e_1) = s(e_n)$ and we say that $s(e_1)$ is the *base point of* α. A *simple loop* in E is a loop $\alpha = e_1 \cdots e_n$ such that $s(e_i) \neq s(e_j)$ for $i \neq j$. We say that E satisfies *Condition (K)* if every vertex is either the base point of at least two simple loops or is not the base point of a loop. It is well-known that if A is a Cuntz-Krieger algebra, then A is isomorphic to $C^*(E)$, where E is a finite graph with no sinks. If, in addition, A is purely infinite, then E will also satisfy Condition (K).

Proposition 6.2 *Let E be a graph with finitely many vertices. Suppose there exists a vertex w in E such that*

(i) $\{w\}$ is a hereditary and saturated subset of E^0;

(ii) $|\{e \in E^1 : s(e) = w\}|$ is either equal to 0 or ∞;

(iii) *for every $v \in E^0 \setminus \{w\}$, there are finitely many edges from v to w and there exists at least one $v \in E^0 \setminus \{w\}$ such that there exists an edge from v to w; and*

(iv) *every vertex $v \in E^0 \setminus \{w\}$ emits finitely many edges and is the base point of at least two loops of length one.*

Then for every full projection $p \in C^*(E) \otimes \mathcal{K}$, we have that $p(C^*(E) \otimes \mathcal{K})p$ is isomorphic to a graph algebra. Consequently, if A is a unital C^*-algebra such that $A \otimes \mathcal{K} \cong C^*(E) \otimes \mathcal{K}$, then A is isomorphic to a graph algebra.

Proof: Let $\{e_{ij}\}$ be a system of matrix units for $\mathcal{K}$. Throughout the proof, if p is a projection in $C^*(E)$ and $n \in \mathbb{N}$, then set $np = \overbrace{p \oplus \cdots \oplus p}^{n}$ in $C^*(E) \otimes \mathcal{K}$. Let $\{p_v, s_e : v \in E^0, e \in E^1\}$ be a Cuntz-Krieger E-family generating $C^*(E)$. Since the only vertex in E that is a singular vertex, i.e., emits no edges or infinitely many edges, is w, by [HL+14, Theorem 3.4 and Corollary 3.5],

$$p \sim \left(\bigoplus_{v \in S} n_v p_v\right) \oplus n_1 \left(p_w - \sum_{e \in T_1} s_e s_e^*\right) \oplus \cdots \oplus n_k \left(p_w - \sum_{e \in T_k} s_e s_e^*\right), \quad (7)$$

where $n_v > 0$ for all $v \in S$, $n_i \geq 0$ for all i, $S \subseteq E^0 \setminus \{w\}$, and T_i is a finite (possibly empty) subset of $s^{-1}(w)$ for all i. Arguing as in [AR15, Lemma 4.6], we have that the projection on the right hand side of (7) is Murray-von Neumann equivalent to

$$q = \bigoplus_{v \in E^0} m_v p_v$$

where $m_v > 0$ for all $v \in E^0$. We use the fact that if S is a finite subset of $s^{-1}(w)$, then

$$|S| p_w \oplus \left(p_w - \sum_{e \in S} s_e s_e^*\right) \sim \left(\sum_{e \in S} s_e s_e^*\right) \oplus \left(p_w - \sum_{e \in S} s_e s_e^*\right) \sim p_w$$

and the fact that if $v_0 \in E^0 \setminus \{w\}$ with $s^{-1}(v_0) \cap r^{-1}(w) \neq \emptyset$, then for any n, we have that $p_{v_0} \sim np_w \oplus p_{v_0} \oplus \left(\bigoplus_{v \in E^0} m'_v p_v\right)$ for $m'_v \geq 0$. Now, arguing as in [AR15, Proposition 4.7], we have that $q(C^*(E) \otimes \mathcal{K})q$ is isomorphic to a graph algebra. Since $p \sim q$, we have that $p(C^*(E) \otimes \mathcal{K})p \cong q(C^*(E) \otimes \mathcal{K})q$. Therefore, $p(C^*(E) \otimes \mathcal{K})p$ is isomorphic to a graph algebra.

For the last part of the proposition note that $A \cong p(C^*(E) \otimes \mathcal{K})p$, where p is the projection given by the image of $1_A \otimes e_{11}$ under some isomorphism from $A \otimes \mathcal{K}$ to $C^*(E) \otimes \mathcal{K}$. Since $1_A \otimes e_{11}$ is full in $A \otimes \mathcal{K}$, we have that p is full in $C^*(E) \otimes \mathcal{K}$. □

6.3 Semigroup C*-algebras and graph algebras

We now determine when a C*-algebra associated to an Artin monoid is isomorphic to a graph algebra. To do this, we need to determine when an extension of two graph algebras is isomorphic to a graph algebra. In spite of substantial effort the extension problem for graph algebras has not be completely resolved even for the single non-trivial ideal case. Moreover, the results in the literature are not sufficient for our purposes. The following ad hoc result will give us what we need.

Lemma 6.3 *For each i, let A_i be a separable, nuclear C*-algebra with an essential ideal I_i such that I_i is isomorphic to either $\mathcal{K}$ or a purely infinite simple C*-algebra with trivial K_1 group, A_i/I_i satisfies the Universal Coefficient Theorem, and $K_1(A_i/I_i) = \{0\}$ or $K_0(A_i/I_i)$ is a free group (possibly $K_0(A_i/I_i) = \{0\}$). Suppose there exist isomorphisms $\beta\colon I_1 \otimes \mathcal{K} \to I_2 \otimes \mathcal{K}$, $\alpha\colon (A_1/I_1) \otimes \mathcal{K} \to (A_2/I_2) \otimes \mathcal{K}$, and $\eta_*\colon K_*(A_1 \otimes \mathcal{K}) \to K_*(A_2 \otimes \mathcal{K})$ such that $(K_*(\beta), \eta_*, K_*(\alpha))\colon K_{\mathrm{six}}(A_1 \otimes \mathcal{K}; I \otimes \mathcal{K}) \to K_{\mathrm{six}}(A_2 \otimes \mathcal{K}; I_2 \otimes \mathcal{K})$ is an isomorphism. Then $A_1 \otimes \mathcal{K} \cong A_2 \otimes \mathcal{K}$.*

Proof: Let $e_1 : 0 \to I_2 \otimes \mathcal{K} \to B_1 \to (A_1/I_1) \otimes \mathcal{K} \to 0$ be the extension obtained by pushing forward the extension $0 \to I_1 \otimes \mathcal{K} \to A_1 \otimes \mathcal{K} \to (A_1/I_1) \otimes \mathcal{K} \to 0$ via the isomorphism β and let $e_2 : 0 \to I_2 \otimes \mathcal{K} \to B_2 \to (A_1/I_1) \otimes \mathcal{K} \to 0$ be the extension obtained by pulling back the extension $0 \to I_2 \otimes \mathcal{K} \to A_2 \otimes \mathcal{K} \to (A_2/I_2) \otimes \mathcal{K} \to 0$ via the isomorphism α. Note that there exist isomorphisms $\phi_1\colon A_1 \otimes \mathcal{K} \to B_1$ and $\phi_2\colon B_2 \to A_2 \otimes \mathcal{K}$ such that $(K_*(\beta), K_*(\phi_1), K_*(\mathrm{id}_{(A_1/I_1)\otimes\mathcal{K}}))\colon K_{\mathrm{six}}(A_1 \otimes \mathcal{K}; I_1 \otimes \mathcal{K}) \to K_{\mathrm{six}}(B_1; I_2 \otimes \mathcal{K})$ and $(K_*(\mathrm{id}_{I_2\otimes\mathcal{K}}), K_*(\phi_2), K_*(\alpha))\colon K_{\mathrm{six}}(B_2; I_2 \otimes \mathcal{K}) \to K_{\mathrm{six}}(A_2 \otimes \mathcal{K}; I_2 \otimes \mathcal{K})$ are isomorphisms. Then $(K_*(\mathrm{id}_{I_2\otimes\mathcal{K}}), K_*(\phi_2^{-1}) \circ \eta_* \circ K_*(\phi_1^{-1}), K_*(\mathrm{id}_{(A_1/I_1)\otimes\mathcal{K}}))$ is an isomorphism from $K_{\mathrm{six}}(B_1; I_2 \otimes \mathcal{K})$ to $K_{\mathrm{six}}(B_2; I_2 \otimes \mathcal{K})$.

We claim that $[\tau_{e_1}] = [\tau_{e_2}]$ in $\mathrm{Ext}((A_1/I_1)\otimes\mathcal{K}, I_2\otimes\mathcal{K})$. Since A_1/I_1 satisfies the Universal Coefficient Theorem, we may identify $\mathrm{Ext}((A_1/I_1) \otimes \mathcal{K}, I_2 \otimes \mathcal{K})$ with $\mathrm{KK}^1((A_1/I_1)\otimes\mathcal{K}, I_2\otimes\mathcal{K})$. Note that $\mathrm{Ext}^1_{\mathbb{Z}}(K_1((A_1/I_1)\otimes\mathcal{K}), K_1(I_2\otimes\mathcal{K})) = \{0\}$ since $K_1(I_2) = \{0\}$. Suppose $K_1(A_1/I_1) = \{0\}$. Then $K_1(\tau_{e_i}) = \{0\}$. Since $K_1(I_2) = \{0\}$, we have that $K_0(\tau_{e_i}) = \{0\}$. Hence, $K_*(\tau_{e_i}) = \{0\}$. By the Universal Coefficient Theorem, $[\tau_{e_i}]$ can be identified with the element in $\mathrm{Ext}^1_{\mathbb{Z}}(K_0((A_1/I_1) \otimes \mathcal{K}), K_0(I_2 \otimes \mathcal{K}))$ given by $K_{\mathrm{six}}(B_i; I_2 \otimes \mathcal{K})$. Since $(K_*(\mathrm{id}_{I_2\otimes\mathcal{K}}), K_*(\phi_2^{-1}) \circ \eta_* \circ K_*(\phi_1^{-1}), K_*(\mathrm{id}_{(A_1/I_1)\otimes\mathcal{K}}))$ is an isomorphism from $K_{\mathrm{six}}(B_1; I_2 \otimes \mathcal{K})$ to $K_{\mathrm{six}}(B_2; I_2 \otimes \mathcal{K})$ we have that $K_{\mathrm{six}}(B_1; I_2 \otimes \mathcal{K})$ and $K_{\mathrm{six}}(B_2; I_2\otimes\mathcal{K})$ induce the same element in $\mathrm{Ext}^1_{\mathbb{Z}}(K_0((A_1/I_1)\otimes\mathcal{K}), K_0(I_2\otimes\mathcal{K}))$. Hence, $[\tau_{e_1}] = [\tau_{e_2}]$ in $\mathrm{Ext}((A_1/I_1) \otimes \mathcal{K}, I_2 \otimes \mathcal{K})$. Suppose $K_0(A_1/I_1)$ is a free group (possibly the zero group). By the Universal Coefficient, $[\tau_{e_i}]$ is completely determined by $K_*(\tau_{e_i})$. Since $(K_*(\mathrm{id}_{I_2\otimes\mathcal{K}}), K_*(\phi_2^{-1}) \circ \eta_* \circ K_*(\phi_1^{-1}), K_*(\mathrm{id}_{(A_1/I_1)\otimes\mathcal{K}}))$ is an isomorphism from $K_{\mathrm{six}}(B_1; I_2 \otimes \mathcal{K})$ to $K_{\mathrm{six}}(B_2; I_2 \otimes \mathcal{K})$, we have that $K_*(\tau_{e_1}) = K_*(\tau_{e_2})$. Hence, $[\tau_{e_1}] = [\tau_{e_2}]$ in $\mathrm{Ext}((A_1/I_1) \otimes \mathcal{K}, I_2 \otimes \mathcal{K})$.

In both cases, we have shown that $[\tau_{e_1}] = [\tau_{e_2}]$ in $\mathrm{Ext}((A_1/I_1) \otimes \mathcal{K}, I_2 \otimes \mathcal{K})$, proving our claim. By Proposition 6.1, we have that $B_1 \cong B_2$. Therefore, $A_1 \otimes \mathcal{K} \cong A_2 \otimes \mathcal{K}$. □

Lemma 6.4 *Let A be a unital, separable, nuclear C^*-algebra with an essential ideal I such that $I \cong \mathcal{K}$ or $I \cong \mathcal{O}_\infty \otimes \mathcal{K}$ and A/I is isomorphic to a purely infinite Cuntz-Krieger algebra. If $K_1(A/I) = \{0\}$ or $K_0(A/I)$ is a free group (possibly $K_0(A/I) = \{0\}$), then A is isomorphic to a graph algebra.*

Proof: By [ABK14, Theorem 4.4] and [Res06, Proposition 8.3], there exists a finite graph F such that each vertex of F is the base point of at least two loop of length one and there exists an isomorphism $\phi\colon C^*(F) \otimes \mathcal{K} \to A/I \otimes \mathcal{K}$. Let $\psi\colon C^*(G) \otimes \mathcal{K} \to I$ be an isomorphism such that $K_*(\psi) = \mathrm{id}$, where G is the graph $\{v\}$ with one vertex and no edges if $I \cong \mathcal{K}$ and G is the graph with one vertex $\{v\}$ with infinitely many edges when $I \cong \mathcal{O}_\infty \otimes \mathcal{K}$. By [EKTW, Lemma 5.2(r1) and Proposition 5.5], there exists a graph E with the properties that

(1) $E^0 = G^0 \sqcup F^0$,

(2) E^1 is the union of G^1 and F^1 together with a finite nonzero number of edges from each $w \in F^0$ to v, and

(3) there exist an isomorphism $\alpha_*\colon K_*(C^*(E)) \to K_*(A)$ with the property that $(K_*(\psi), \alpha_*, K_*(\phi))$ is an isomorphism from $K_{\mathrm{six}}(C^*(E); I_{\{v\}})$ to $K_{\mathrm{six}}(A; I)$.

Note that $I_{\{v\}} \otimes \mathcal{K}$ is an essential ideal of $C^*(E) \otimes \mathcal{K}$ and there exist an isomorphism $\overline{\alpha}_*\colon K_*(C^*(E) \otimes \mathcal{K}) \to K_*(A \otimes \mathcal{K})$ such that $(K_*(\psi \otimes \mathrm{id}_{\mathcal{K}}), \overline{\alpha}_*, K_*(\phi \otimes \mathrm{id}_{\mathcal{K}}))$ is an isomorphism from $K_{\mathrm{six}}(C^*(E) \otimes \mathcal{K}; I_{\{v\} \otimes \mathcal{K}})$ to $K_{\mathrm{six}}(A \otimes \mathcal{K}; I \otimes \mathcal{K})$. Also, note that $I \cong I_{\{v\}} = \mathcal{K}$ or $I \cong I_{\{v\}} \cong \mathcal{O}_\infty \otimes \mathcal{K}$. By Lemma 6.3, $A \otimes \mathcal{K} \cong C^*(E) \otimes \mathcal{K}$. Therefore, A is isomorphic to a graph algebra by Proposition 6.2. □

Lemma 6.5 *For each $m \in \mathbb{N}$, for each $n \geq 0$, the smallest nonzero ideal I of $E_m^{\pm 1} \otimes \bigotimes_{k=1}^n E_2^{r_k}$ is isomorphic to $\mathcal{K}$ and $(E_m^{\pm 1} \otimes \bigotimes_{k=1}^n E_2^{r_k})/I$ is isomorphic to a Cuntz-Krieger algebra with vanishing K_1-group.*
Consequently, $E_m^{\pm 1} \otimes \bigotimes_{k=1}^n E_2^{r_k}$ and $E_m^{\pm 1} \otimes \bigotimes_{k=1}^n E_2^{r_k} \otimes \mathcal{O}_\infty$ are isomorphic to graph algebras.

Proof: Note that for each $m \in \mathbb{N}$, by [EKTW, Theorem 7.2], E_m^{+1}, E_m^{-1}, $E_m^{+1} \otimes \mathcal{O}_\infty$, and $E_m^{-1} \otimes \mathcal{O}_\infty$ are graph algebras with $E_m^{\pm 1}/\mathcal{K}$ and $(E_m^{\pm 1} \otimes \mathcal{O}_\infty)/(\mathcal{K} \otimes \mathcal{O}_\infty) \cong (E_m^{\pm 1}/\mathcal{K}) \otimes \mathcal{O}_\infty$ are isomorphic to purely infinite Cuntz-Krieger algebras. Therefore, we may assume that $n \geq 1$.
For notational convenience, set $A = E_m^{\pm 1} \otimes \bigotimes_{k=1}^n E_2^{r_k}$. Note that $I = \bigotimes_{k=1}^{n+1} \mathcal{K}$. Let $J = \mathcal{K} \otimes \bigotimes_{k=1}^n E_2^{r_k}$. Then J is a primitive ideal and $A/J \cong \mathcal{O}_m \otimes \bigotimes_{k=1}^n E_2^{r_k}$. We will show that J/I is stably isomorphic to an $\mathcal{O}_2$-absorbing Cuntz-Krieger

algebra, A/J is isomorphic to a Cuntz-Krieger algebra with vanishing boundary maps, and the boundary maps in K-theory induced by the extension $0 \to J/I \to A/I \to A/J \to 0$ are zero.

We will first prove that J/I is $\mathcal{O}_2$-absorbing. Note that it is enough to show that $(\bigotimes_{k=1}^n E_2^{r_k}) / (\bigotimes_{k=1}^n \mathcal{K})$ is $\mathcal{O}_2$-absorbing since $J/I \cong \mathcal{K} \otimes (\bigotimes_{k=1}^n E_2^{r_k}) / (\bigotimes_{k=1}^n \mathcal{K})$. Since $E_2^{\pm}/\mathcal{K} \cong \mathcal{O}_2$ which is $\mathcal{O}_2$-absorbing by [KP00, Theorem 3.8], we have that $(\bigotimes_{k=1}^n E_2^{r_k}) / (\bigotimes_{k=1}^n \mathcal{K})$ is $\mathcal{O}_2$-absorbing for $n = 1$. Suppose $(\bigotimes_{k=1}^m E_2^{r_k}) / (\bigotimes_{k=1}^m \mathcal{K})$ is $\mathcal{O}_2$-absorbing for $1 \leq m < n$. Consider the extension

$$
\begin{aligned}
0 \to (E_2^{r_1} \otimes \textstyle\bigotimes_{k=2}^n \mathcal{K}) / (\bigotimes_{k=1}^n \mathcal{K}) \to \\
\to (\textstyle\bigotimes_{k=1}^n E_2^{r_k}) / (\bigotimes_{k=1}^n \mathcal{K}) \to (\bigotimes_{k=1}^n E_2^{r_k}) / (E_2^{r_1} \otimes \bigotimes_{k=2}^n \mathcal{K}) \to 0
\end{aligned}
$$

Now, $(E_2^{r_1} \otimes \bigotimes_{k=2}^n \mathcal{K})/(\otimes_{k=1}^n \mathcal{K}) \cong (E_2^{r_1}/\mathcal{K}) \otimes \bigotimes_{k=2}^n \mathcal{K} \cong \mathcal{O}_2 \otimes \bigotimes_{k=2}^n \mathcal{K}$ which is $\mathcal{O}_2$-absorbing by [KP00, Theorem 3.8]. Since $(\bigotimes_{k=1}^n E_2^{r_k})/(E_2^{r_1} \otimes \bigotimes_{k=2}^n \mathcal{K}) \cong E_2^{r_1} \otimes ((\bigotimes_{k=2}^n E_2^{r_k}) / (\bigotimes_{k=2}^n \mathcal{K}))$ and because of the inductive hypothesis, we have that $(\otimes_{k=1}^n E_2^{r_k})/(E_2^{r_1} \otimes \bigotimes_{k=2}^n \mathcal{K})$ is $\mathcal{O}_2$-absorbing. Hence, by [KP00, Theorem 3.8] and [TW07, Corollary 4.3], $(\bigotimes_{k=1}^n E_2^{r_k}) / (\bigotimes_{k=1}^n \mathcal{K})$ is $\mathcal{O}_2$-absorbing. This proves our claim.

Since J/I is $\mathcal{O}_2$-absorbing and J/I has finitely many ideals, by [Kir00], J/I is stably isomorphic to a Cuntz-Krieger algebra with vanishing boundary maps. This is because for any finite T_0-space X, there exists an $\mathcal{O}_2$-absorbing Cuntz-Krieger algebra with primitive ideal space X. We also note that the boundary maps in K-theory induced by the extension $0 \to J/I \to A/I \to A/J \to 0$ are zero since $K_*(J/I) = \{0\}$.

We now show that A/J is isomorphic to a Cuntz-Krieger algebra with vanishing boundary maps. Recall that $A/J \cong \mathcal{O}_m \otimes \bigotimes_{k=1}^n E_2^{r_k}$. Hence, every simple sub-quotient of A/J is isomorphic to $\mathcal{O}_m \otimes (\mathcal{I}_2/\mathcal{I}_1)$ where $\mathcal{I}_1, \mathcal{I}_2$ are ideals of $\bigotimes_{k=1}^n E_2^{r_k}$ with $\mathcal{I}_1 \subseteq \mathcal{I}_2$ and $\mathcal{I}_2/\mathcal{I}_1$ simple. Note that if $\mathcal{I}_1, \mathcal{I}_2$ are ideals of $\bigotimes_{k=1}^n E_2^{r_k}$ with $\mathcal{I}_1 \subseteq \mathcal{I}_2$ and $\mathcal{I}_2/\mathcal{I}_1$ simple, then $\mathcal{I}_2/\mathcal{I}_1 \cong \bigotimes_{k=1}^n B_k$ where B_k is a simple sub-quotient of $E_2^{r_k}$. Hence, every simple sub-quotient of $\bigotimes_{k=1}^n E_2^{r_k}$ is either isomorphic to $\bigotimes_{k=1}^n \mathcal{K}$ or is $\mathcal{O}_2$-absorbing. Hence, every simple sub-quotient of $\mathcal{O}_m \otimes \bigotimes_{k=1}^n E_2^{r_k}$ is either stably isomorphic to $\mathcal{O}_m$ or $\mathcal{O}_2$. So every simple sub-quotient of A/J is stably isomorphic to a Cuntz-Krieger algebra. Consider the extension $e : 0 \to \mathcal{O}_m \otimes \mathcal{I}_1 \to \mathcal{O}_m \otimes \mathcal{I}_2 \to \mathcal{O}_m \otimes (\mathcal{I}_2/\mathcal{I}_1) \to 0$ with $\mathcal{I}_2/\mathcal{I}_1$ simple. If $\mathcal{I}_1 = \{0\}$ and $\mathcal{I}_2 = \bigotimes_{k=1}^n \mathcal{K}$, then $\mathcal{O}_m \otimes \mathcal{I}_1 = \{0\}$ which implies that e has vanishing boundary maps. If $\mathcal{I}_2/\mathcal{I}_1$ is $\mathcal{O}_2$-absorbing, then $K_*(\mathcal{O}_m \otimes (\mathcal{I}_2/\mathcal{I}_1)) = \{0\}$ which also implies that e has vanishing boundary maps. By [Ben14, Corollary 3.6], we have that $A/J \cong \mathcal{O}_m \otimes \bigotimes_{k=1}^n E_2^{r_k}$ has vanishing boundary maps. Therefore, by [Ben14, Corollary 8.2], A/J is isomorphic to a Cuntz-Krieger algebra with vanishing boundary maps. This finishes the proof of the above claim.

The above claim shows that all the assumptions in [Ben14, Proposition 3.7, Proposition 3.10, and Corollary 8.4] are satisfied. Thus, A/I is isomorphic to a purely infinite Cuntz-Krieger algebra.

We now show that $K_1((E_m^{\pm 1} \otimes \bigotimes_{k=1}^n E_2^{r_k})/I) \cong \{0\}$. Since $0 \to J/I \to A/I \to A/J \to 0$ has vanishing boundary maps and the fact that J/I is $\mathcal{O}_2$-absorbing, we have that the surjective map $A/I \to A/J$ induces an injective map $K_1(A/I) \to K_1(A/J)$. Since every simple sub-quotient of A/J is stably isomorphic to $\mathcal{O}_m$ or $\mathcal{O}_2$ and since A/J has finitely many ideals, one can show that $K_1(A/J) = \{0\}$. Therefore, $K_1(A/I) = \{0\}$.
Lemma 6.4 implies that $E_m^{\pm 1} \otimes \bigotimes_{k=1}^n E_{m_k}^{r_k}$ and $E_m^{\pm 1} \otimes \bigotimes_{k=1}^n E_{m_k}^{r_k} \otimes \mathcal{O}_\infty$ are isomorphic to graph algebras. □

Lemma 6.6 *Let $m_1, m_2, \dots m_n \in \mathbb{N}$. Then*

(1) $\bigotimes_{k=1}^n E_{m_k}^{\pm 1}$ is stably isomorphic to unital graph algebra if and only if whenever there exists an i such that $m_i \in \{1\} \sqcup \mathbb{Z}_{\geq 3}$, we have that $m_j = 2$ for all $j \neq i$.

(2) $\bigotimes_{k=1}^n E_{m_k}^{\pm 1} \otimes \mathcal{O}_\infty$ is stably isomorphic to unital graph algebra if and only if whenever there exists an i such that $m_i \in \{1\} \sqcup \mathbb{Z}_{\geq 3}$, we have that $m_j = 2$ for all $j \neq i$.

Proof: We prove (1). (2) is proved in a similar way.
Suppose whenever there exists an i such that $m_i \in \{1\} \sqcup \mathbb{Z}_{\geq 3}$, we have that $m_j = 2$ for all $j \neq i$. By Lemma 6.5, $\bigotimes_{k=1}^n E_{m_k}^{\pm 1}$ is isomorphic to a graph algebra. So, also stably isomorphic to a unital graph algebra.
Suppose $\bigotimes_{k=1}^n E_{m_k}^{\pm 1}$ is stably isomorphic to graph algebra. Note that $E_m^{\pm 1} \otimes \mathcal{K} \cong E_m^{+1} \otimes \mathcal{K}$ for any m. Therefore, it is enough to prove the case $\bigotimes_{k=1}^n E_{m_k}^{+1}$. Note that $\bigotimes_{k=1}^n E_{m_k}^{+1}$ has finitely many ideals. Since $\bigotimes_{k=1}^n E_{m_k}^{+1}$ is stably isomorphic to a unital graph algebra $C^*(E)$, we have that $C^*(E)$ has finitely many ideals. Therefore, every sub-quotient of $C^*(E)$ is stably isomorphic to a unital graph algebra with finitely many ideals. Consequently, every sub-quotient of $\bigotimes_{k=1}^n E_{m_k}^{+1}$ is stably isomorphic to a unital graph algebra with finitely many ideals.
Suppose there exists i and j such that $m_i, m_j \in \{1\} \sqcup \mathbb{Z}_{\geq 3}$. Let $I = \bigotimes_{k=1}^n I_k$ be the ideal of $\bigotimes_{k=1}^n E_{m_k}^{+1}$ where $I_k = \mathcal{K}$ if $k \notin \{i, j\}$, $I_i = E_{m_i}^{+1}$, and $I_j = E_{m_j}^{+1}$. From the above observation we must have that every sub-quotient of I is stably isomorphic to a unital graph algebra with finitely many ideals. Note that I is stably isomorphic to $E_{m_i}^{+1} \otimes E_{m_j}^{+1}$ and $E_{m_i}^{+1} \otimes E_{m_j}^{+1}$ has a quotient isomorphic to $\mathcal{O}_{m_i} \otimes \mathcal{O}_{m_j}$. Therefore, $\mathcal{O}_{m_i} \otimes \mathcal{O}_{m_j}$ is stably isomorphic to a graph algebra.
Let $\mathcal{K} \otimes \mathcal{K}$ be the smallest non-zero ideal of $E_{m_i}^{+1} \otimes E_{m_j}^{+1}$. By the Künneth formula, $K_0(E_{m_i}^{+1} \otimes E_{m_j}^{+1}) \cong \mathbb{Z}$ and $K_1(E_{m_i}^{+1} \otimes E_{m_j}^{+1}) = \{0\}$, and hence the extension $0 \to \mathcal{K} \otimes \mathcal{K} \to E_{m_i}^{+1} \otimes E_{m_j}^{+1} \to (E_{m_i}^{+1} \otimes E_{m_j}^{+1})/(\mathcal{K} \otimes \mathcal{K}) \to 0$ induces a six-term exact sequence in K-theory of the form

$$\begin{tikzcd}
\mathbb{Z} \arrow[r] & \mathbb{Z} \arrow[r] & K_0((E_{m_i}^{+1} \otimes E_{m_j}^{+1})/(\mathcal{K} \otimes \mathcal{K})) \arrow[d] \\
K_1((E_{m_i}^{+1} \otimes E_{m_j}^{+1})/(\mathcal{K} \otimes \mathcal{K})) \arrow[u] & 0 \arrow[l] & 0. \arrow[l]
\end{tikzcd}$$

In particular, $K_0((E_{m_i}^{+1} \otimes E_{m_j}^{+1})/(\mathcal{K} \otimes \mathcal{K}))$ and $K_1((E_{m_i}^{+1} \otimes E_{m_j}^{+1})/(\mathcal{K} \otimes \mathcal{K}))$ are cyclic groups.
Since $(E_{m_i}^{+1} \otimes E_{m_j}^{+1})/(\mathcal{K} \otimes \mathcal{K})$ is stably isomorphic to a graph algebra with finitely many ideals, $(E_{m_i}^{+1} \otimes E_{m_j}^{+1})/(\mathcal{K} \otimes \mathcal{K})$ has real rank zero. Therefore, the quotient of $(E_{m_i}^{+1} \otimes E_{m_j}^{+1})/(\mathcal{K} \otimes \mathcal{K})$ by the ideal $(\mathcal{K} \otimes E_{m_j}^{+1} + E_{m_i}^{+1} \otimes \mathcal{K})/(\mathcal{K} \otimes \mathcal{K})$ induces the following six-term exact sequence

$$\begin{tikzcd} K_0(\mathcal{O}_{m_i}) \oplus K_0(\mathcal{O}_{m_j}) \arrow[r] & K_0((E_{m_i}^{+1} \otimes E_{m_j}^{+1})/(\mathcal{K} \otimes \mathcal{K})) \arrow[r] & K_0(\mathcal{O}_{m_i} \otimes \mathcal{O}_{m_j}) \arrow[d, "0"] \\ K_1(\mathcal{O}_{m_i} \otimes \mathcal{O}_{m_j}) \arrow[u] & K_1((E_{m_i}^{+1} \otimes E_{m_j}^{+1})/(\mathcal{K} \otimes \mathcal{K})) \arrow[l] & K_1(\mathcal{O}_{m_i}) \oplus K_1(\mathcal{O}_{m_j}). \arrow[l] \end{tikzcd} \tag{8}$$

Using the Künneth formula, we get

$$K_0(\mathcal{O}_{m_i} \otimes \mathcal{O}_{m_j}) = K_1(\mathcal{O}_{m_i} \otimes \mathcal{O}_{m_j}) = \begin{cases} \mathbb{Z}_{\gcd(m_i-1, m_j-1)} & \text{if } m_i, m_j \geq 3 \\ K_1(\mathcal{O}_{m_i}) \oplus K_0(\mathcal{O}_{m_i}) & \text{if } m_j = 1 \\ K_1(\mathcal{O}_{m_j}) \oplus K_0(\mathcal{O}_{m_j}) & \text{if } m_i = 1. \end{cases}$$

Since $\mathcal{O}_{m_i} \otimes \mathcal{O}_{m_j}$ is stably isomorphic to a unital graph algebra, $\gcd(m_i - 1, m_j - 1) = 1$ if $m_i, m_j \geq 3$ and $m_i = 1$ if and only if $m_j = 1$.
Suppose $m_i, m_j \geq 3$. The exactness of Diagram (8) implies that $K_0((E_{m_i}^{+1} \otimes E_{m_j}^{+1})/(\mathcal{K} \otimes \mathcal{K})) \cong K_0(\mathcal{O}_{m_i}) \oplus K_0(\mathcal{O}_{m_j}) \cong \mathbb{Z}_{m_i-1} \oplus \mathbb{Z}_{m_j-1}$ which contradicts the fact that $K_0(E_{m_i}^{+1} \otimes E_{m_j}^{+1})/(\mathcal{K} \otimes \mathcal{K}))$ is a cyclic group.
Suppose $m_i = 1$. Then $m_j = 1$. Then by the exactness of Diagram (8), $K_1((E_{m_i}^{+1} \otimes E_{m_j}^{+1})/(\mathcal{K} \otimes \mathcal{K}))$ has a sub-group isomorphic to $K_1(\mathcal{O}_{m_i}) \oplus K_1(\mathcal{O}_{m_j}) \cong \mathbb{Z} \oplus \mathbb{Z}$. This can not happen since $K_1((E_{m_i}^{+1} \otimes E_{m_j}^{+1})/(\mathcal{K} \otimes \mathcal{K}))$ is a cyclic group.
□
Let the notation be as in Definition 4.1.

THEOREM 6.7 *Let Γ be a countable graph. Then $C^*(A_\Gamma^+)$ is isomorphic to a graph algebra if and only if one of the following holds*

1. $t(\Gamma) = 1$, $o(\Gamma) = 0$ *and* $N_k(\Gamma) = 0$ *for all* k
2. $t(\Gamma) = 0$, $N_{-1}(\Gamma) + N_1(\Gamma) < \infty$ *and*
$$\sum_{|k| \neq 1} N_k(\Gamma) \leq 1$$

Proof: Suppose there exists an isomorphism $\psi \colon C^*(A_\Gamma^+) \to C^*(E)$ for some countable directed graph E. Since $C^*(A_\Gamma^+)$ is unital, $C^*(E)$ is unital. Let $\Gamma_i = (V_i, E_i)$ be the co-irreducible components of Γ. To prove (1), let I be the ideal of $C^*(A_\Gamma^+)$ generated by $\{\bigotimes_j J_{ij}\}_i$ where $J_{ij} = C^*(A_{\Gamma_j}^+)$ if $j \neq i$ and

$$J_{ii} = \begin{cases} \mathcal{K} & \text{if } 1 < |V_i| < \infty \\ 0 & \text{otherwise.} \end{cases}$$

Then $C^*(A_\Gamma^+)/I \cong \bigotimes_i C^*(A_{\Gamma_i}^+)/J_{ii}$ where $C^*(A_{\Gamma_i}^+)/J_{ii}$ is a Kirchberg algebra if $|V_i| \geq 2$ and $C^*(A_{\Gamma_i}^+)/J_{ii} \cong \mathcal{T}$ otherwise. In particular,

$$\mathrm{Prim}(C^*(A_\Gamma^+)/I) \cong \begin{cases} \prod_{k=1}^{t(\Gamma)} \mathrm{Prim}(\mathcal{T}) & \text{if there exists } i \text{ with } |V_i| = 1 \\ \{\bullet\} & \text{otherwise.} \end{cases}$$

Note that I is generated by projections. Therefore, $\psi(I)$ is generated by projections and hence is a gauge-invariant ideal of $C^*(E)$. Hence, by [BPRS00, Corollary 3.5 and Theorem 3.6], $C^*(E)/\psi(I)$ is isomorphic to a graph algebra. Since $C^*(A_\Gamma^+)/I \cong C^*(E)/\psi(I)$, we have that $C^*(A_\Gamma^+)/I$ is isomorphic to a unital graph algebra. Note that $C^*(A_\Gamma^+)/I$ is $\mathcal{O}_\infty$-absorbing (if there exists i such that $|V_i| \geq 2$) or $C^*(A_\Gamma^+)/I \cong \bigotimes_{k=1}^{t(\Gamma)} \mathcal{T}$.

Suppose $C^*(A_\Gamma^+)/I$ is $\mathcal{O}_\infty$-absorbing. Since any unital $\mathcal{O}_\infty$-absorbing graph algebra has a finite primitive ideal space, we must have that $t(\Gamma) = 0$. Suppose $C^*(A_\Gamma^+)/I$ is not $\mathcal{O}_\infty$-absorbing. Then $C^*(A_\Gamma^+)/I \cong \bigotimes_{k=1}^{t(\Gamma)} \mathcal{T}$. Let J be the ideal generated by $\{\bigotimes_j J_{ij}\}_i$ where $J_{ij} = \mathcal{T}$ if $j \neq i$ and $J_{ii} = \mathcal{K}$, then J is an ideal generated by projections such that $\left(\bigotimes_{k=1}^{t(\Gamma)} \mathcal{T}\right)/J \cong C(\mathbb{T}^{t(\Gamma)})$. Since $C^*(A_\Gamma^+)/I$ is isomorphic to a graph algebra and every ideal generated by projections in a graph algebra is gauge invariant, by [BPRS00, Corollary 3.5 and Theorem 3.6] every quotient of $C^*(A_\Gamma^+)/I$ by an ideal generated by projections is isomorphic to a graph algebra. Hence, $C(\mathbb{T}^{t(\Gamma)}) \cong \left(\bigotimes_{k=1}^{t(\Gamma)} \mathcal{T}\right)/J$ is isomorphic to a unital graph algebra. Since the only unital commutative graph algebra is isomorphic to finite direct sums of $\mathbb{C}$ and $\mathbb{T}$, we must have that $t(\Gamma) = 1$.

In both cases, we have shown that $t(\Gamma) \leq 1$. Suppose $o(\Gamma) \neq 0$ or $N_k(\Gamma) \neq 0$ for some k, then there exists an i such that $C^*(A_{\Gamma_i}^+)/J_{ii}$ is a Kirchberg algebra. Hence, by [KP00, Theorem 3.15] and [TW07, Corollary 3.4] $C^*(A_\Gamma^+)/I \cong \bigotimes_i C^*(A_{\Gamma_i}^+)/J_{ii}$ is an $\mathcal{O}_\infty$-absorbing C*-algebra. Since every unital graph algebra that is $\mathcal{O}_\infty$-absorbing must have finitely many ideals and since

$$\mathrm{Prim}(C^*(A_\Gamma^+)/I) \cong \begin{cases} \prod_{k=1}^{t(\Gamma)} \mathrm{Prim}(\mathcal{T}) & \text{if there exists } i \text{ with } |V_i| = 1 \\ \{\bullet\} & \text{otherwise,} \end{cases}$$

we have that $t(\Gamma) = 0$. Hence, we only get a graph algebra in the case $t(\Gamma) = 1$ when all other data vanish.

Suppose $t(\Gamma) = 0$. Note that $1 < |V_i|$ for all i. Thus, $C^*(A_{\Gamma_i}^+)$ is a unital properly infinite C*-algebra, and $\mathrm{Prim}(C^*(A_{\Gamma_i}^+)) = \{x_i, y_i\}$ with open sets $\{\emptyset, \{x_i\}, \{x_i, y_i\}\}$ when $|V_i| < \infty$ and $\mathrm{Prim}(C^*(A_{\Gamma_i}^+)) \cong \{\bullet\}$ when $|V_i| = \infty$.

We claim that $|\{k\colon N_k(\Gamma) \neq 0\}| < \infty$ and $N_k(\Gamma) < \infty$ for all k. Suppose first $|\{k\colon N_k(\Gamma) \neq 0\}| = \infty$ or $N_k(\Gamma) = \infty$ for some k. Then $C^*(A_\Gamma^+) \cong \bigotimes_{i=1}^{\infty} C^*(A_{\Gamma_i}^+)$ and $C^*(A_\Gamma^+)$ has infinitely many ideals. By Lemma 4.7 , $C^*(A_\Gamma^+)$ is $\mathcal{O}_\infty$-absorbing. Again, using the fact that a unital graph algebra that is

$\mathcal{O}_\infty$-absorbing has finitely many ideals, we have a contradiction. Therefore, $|\{k\colon N_k(\Gamma) \neq 0\}| < \infty$ and $N_k(\Gamma) < \infty$ for all k, proving the claims in (2). Note that

$$C^*(A_\Gamma^+) \cong (E_1^0)^{\otimes N_0(\Gamma)} \otimes \bigotimes_{n=1}^{\infty} (E_{1+n}^{-1})^{\otimes N_{-n}(\Gamma)} \otimes \bigotimes_{n=1}^{\infty} (E_{1+n}^{+1})^{\otimes N_n(\Gamma)} \otimes (\mathcal{O}_\infty)^{\otimes o(\Gamma)}$$

By Lemma 6.6, (1) and (2) hold.
In the other direction, we have in the case (1) that $C^*(A_\Gamma^+) \cong \mathcal{T}$ which is isomorphic to a graph algebra. And in case (2) we have that either

$$C^*(A_\Gamma^+) \cong (E_2^{-1})^{\otimes N_{-1}(\Gamma)} \otimes (E_2^{+1})^{\otimes N_1(\Gamma)} \otimes (\mathcal{O}_\infty)^{\otimes o(\Gamma)},$$

$$C^*(A_\Gamma^+) \cong E_m^{+1} \otimes (E_2^{-1})^{\otimes N_{-1}(\Gamma)} \otimes (E_2^{+1})^{\otimes N_1(\Gamma)} \otimes (\mathcal{O}_\infty)^{\otimes o(\Gamma)}$$

for some $m \neq 2$, or

$$C^*(A_\Gamma^+) \cong E_m^{-1} \otimes (E_2^{-1})^{\otimes N_{-1}(\Gamma)} \otimes (E_2^{+1})^{\otimes N_1(\Gamma)} \otimes (\mathcal{O}_\infty)^{\otimes o(\Gamma)}.$$

for some $m \neq 2$. If $o(\Gamma) \geq 1$, then by [KP00, Theorem 3.15], $(\mathcal{O}_\infty)^{\otimes o(\Gamma)} \cong \mathcal{O}_\infty$. Hence, by Lemma 6.5, $C^*(A_\Gamma^+)$ is isomorphic to a graph algebra. □

Remark 6.8 *The relation between a (undirected, loop-free) graph Γ and a directed graph G_Γ with $C^*(A_\Gamma^+) \cong C^*(G_\Gamma)$ is somewhat opaque, although the proof given above in principle is constructive. In Figure 2 we present eight graphs presenting the C*-algebras given by five-vertex graphs of Figure 1 in the unshaded regions.*

We conclude by establishing semiprojectivity and non-semiprojectivity of $C^*(A_\Gamma^+)$ in a number of cases, covering for instance all graphs with 5 or fewer vertices. We note, however, that this theorem does not contain a full answer to the question of which of the C*-algebras under study are semiprojective. The most basic open case has $N_{-2} = 2$ and may be represented by a graph with 6 vertices.

Theorem 6.9

1. *When $t(\Gamma) > 1$, $C^*(A_\Gamma^+)$ is not semiprojective.*
2. *When $t(\Gamma) = 1$, $C^*(A_\Gamma^+)$ is semiprojective if and only if*
$$o(\Gamma) = \sum_k N_k(\Gamma) = 0.$$
3. *When $t(\Gamma) = 0$, $C^*(A_\Gamma^+)$ is semiprojective when $N_{-1}(\Gamma) + N_1(\Gamma) < \infty$ and*
$$\sum_{|k|\neq 1} N_k(\Gamma) \leq 1$$

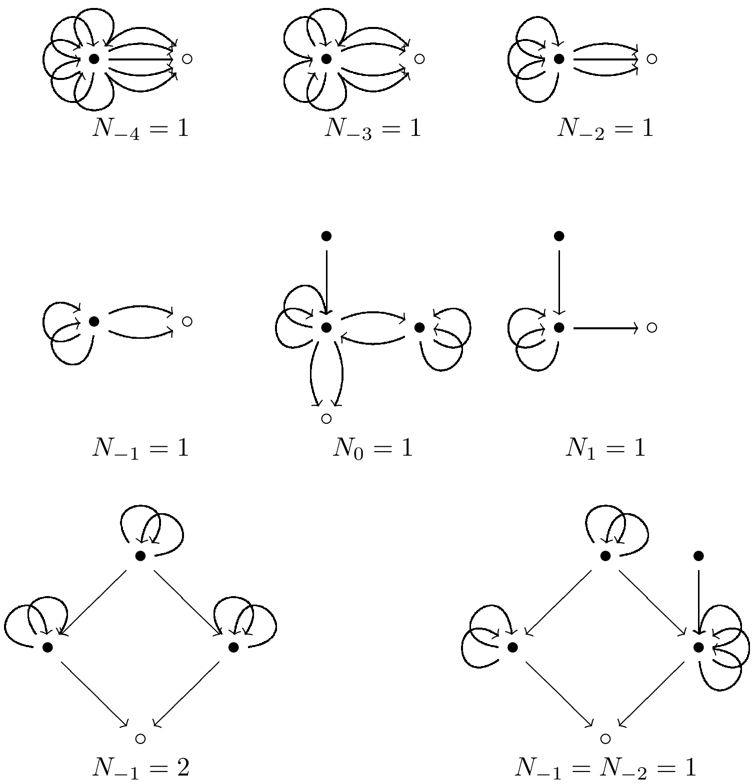


Figure 2: Graphs representing cases from Figure 1

Proof: We first note that by [End13, Corollary 4.4.16], a C*-algebra of the form $A \otimes \mathcal{T}$ with A unital, nuclear, infinite-dimensional and in the UCT-class can never be semiprojective. This proves (1) and (2) since $\mathcal{T}$ itself is trivially semiprojective.
For (3), we first apply Theorem 6.7 to see that $C^*(A_\Gamma^+)$ in this case is a unital graph algebra. We have seen that when $o(\Gamma) > 0$, $C^*(A_\Gamma^+)$ is strongly purely infinite, and when $o(\Gamma) = 0$, there is a minimal ideal $\mathcal{K}$ in $C^*(A_\Gamma^+)$ so that $C^*(A_\Gamma^+)/\mathcal{K}$ is strongly purely infinite. In either case, [EK] applies to guarantee that the C*-algebra is semiprojective. $\square$

References

[AD] Y. Antolin and D. Dreesen, *The Haagerup property is stable under graph products*, arXiv:1305.6748v2.

[ABK14] S.E. Arklint, R. Bentmann, T. Katsura, *The K-theoretical range of Cuntz-Krieger algebras*, J. Funct. Anal. 266 (2014), no. 8, 5448–5466.

[AGR] S.E. Arklint, J. Gabe, and E. Ruiz, *Hereditary C*-subalgebras of graph C*-algebras*, arXiv:1604.03085.

[AR15] S.E. Arklint and E. Ruiz, *Corners of Cuntz-Krieger algebras*, Trans. Amer. Math. Soc.367 (2015), 7595–7612.

[Arv77] W. Arveson, *Notes on extensions C*-algebras*, Duke Math. J. 44 (1977), 329–355.

[BPRS00] T. Bates, D. Pask, I. Raeburn, and W. Szymanski, *C*-algebras of row-finite graphs*, New York J. Math. 6 (2000), 307–324.

[Ben14] R. Bentmann, *Kirchberg X-algebras with real rank zero and intermediate cancellation*, J. Noncommut. Geom. 8 (2014),1061–1081.

[BK] R. Bentmann and M. Köhler, *Universal coefficient theorems for C*-algebras over finite topological spaces*, arXiv:1101.5702v3.

[Bla77] B. Blackadar, *Infinite tensor products of C*-algebras*, Pacific J. Math. 72 (1977), 313–334.

[Bla85] ———, *Shape theory for C*-algebras*, Math. Scand. 56 (1985), 249–275.

[Bla04] ———, *Semiprojectivity in simple C*-algebras*, Adv. Stud. Pure Math. 38 (2004), 1–17.

[BD96] L.G. Brown and M. Dădărlat, *Extensions of C*-algebras and quasidiagonality*, J. London Math. Soc. (2) 53 (1996), no. 3, 582–600.

[CL02] J. Crisp and M. Laca, *On the Toeplitz algebras of right-angled and finite-type Artin groups*, J. Austral. Math. Soc. 72 (2002), 223–245.

[CL07] ———, *Boundary quotients and ideals of Toeplitz algebras of Artin groups groups*, J. Funct. Anal. 242 (2007), 127–156.

[CEL13] J. Cuntz, S. Echterhoff, and X. Li, *On the K-theory of crossed products by automorphic semigroup actions*, Q. J. Math. 64 (2013), 747–784.

[EK] S. Eilers and T. Katsura, *Semiprojectivity and properly infinite projections in graph C^*-algebras*, arXiv:1512.07277.

[EKTW] S. Eilers, T. Katsura, M. Tomforde, and J. West, *The ranges of K-theoretical invariants for nonsimple graph algebras*, Trans. Amer. Math. Soc., 368 (2016), 3811–3847.

[ELP99] S. Eilers, T. A. Loring, and G. K. Pedersen, *Morphisms of extensions of C^*-algebras: Pushing forward the Busby invariant*, Adv. Math. 147 (1999), 74–109.

[ERR] S. Eilers, G. Restorff, and E. Ruiz, *Strong classification of extensions of classifiable C^*-algebras*, arXiv:1301.7695.

[ERR09] ———, *Classification of extensions of classifiable C^*-algebras*, Adv. Math. 222 (2009), 2153–2172.

[ERR10] ———, *On graph C^*-algebras with a linear ideal lattice*, Bull. Malays. Math. Sci. Soc. 33 (2010), no. 2, 233–241.

[ERR13] ———, *Classifying C^*-algebras with both finite and infinite subquotients*, J. Funct. Anal. 265 (2013), 449–468.

[End13] D. Enders, *On the structure of certain classes of semiprojective C^*-algebras*, Ph.D. thesis, Westfälische Wilhelms-Universität Münster, 2013.

[End15] D. Enders, *Semiprojectivity for Kirchberg algebras*, preprint, arXiv:1507.0609, 2015.

[EL91] R. Exel and T.A. Loring, *Invariants of almost commuting unitaries*, J. Funct. Anal. 95 (1991), 364–376.

[FL07] X. Fang and S. Liu, *Extension algebras of Cuntz algebras*, J. Math. Anal. Appl 329 (2007), 655–663.

[HL+14] D. Hay, M. Loving, M. Montgomery, E. Ruiz, and K. Todd, *Nonstable K-theory for Leavitt path algebras*, Rocky Mountain J. Math. 44 (2014), 1817-1850.

[Iva10] N. Ivanov, *The K-theory of Toeplitz C^*-algebras of right-angled Artin groups*, Trans. Amer. Math. Soc. 362 (2010), no. 11, 6003–6027.

[JT91] K.K. Jensen and K. Thomsen, *Elements of KK-theory*, Birkhäuser, Boston, 1991.

[Kir] E. Kirchberg, *The classification of purely infinite C^*-algebras using Kasparov's theory* 3rd draft. Preprint.

[Kir00] E. Kirchberg, *Das nicht-kommutative Michael-Auswahlprinzip und die Klassifikation nicht-einfacher Algebren*, C^*-algebras (Münster, 1999), Springer, Berlin, 2000, pp. 92–141.

[KP00] E. Kirchberg and N. C. Phillips, *Embedding of exact C^*-algebras in the Cuntz algebra $\mathcal{O}_2$*, J. Reine Angew. Math. 525 (2000), 17–53.

[KR02] E. Kirchberg and M. Rørdam, *Infinite non-simple C^*-algebras: absorbing the Cuntz algebras $\mathcal{O}_\infty$*, Adv. Math. 167 (2002), 195–264.

[Li12] X. Li, *Semigroup C^*-algebras and amenability of semigroups*, J. Funct. Anal. 262 (2012), 4302–4340.

[Li13] X. Li, *Nuclearity of semigroup C^*-algebras and the connection to amenability*, Adv. Math. 244 (2013), 626–662.

[Lor97] T.A. Loring, *Lifting solutions to perturbing problems in C^*-algebras*, Fields Institute Monographs, vol. 8, American Mathematical Society, Providence, RI, 1997.

[NR97] G. Niblo and L. Reeves, *Groups acting on CAT(0) cube complexes*, Geom. & Top. 1 (1997), 1–7.

[Par02] L. Paris, *Artin monoids inject in their groups*, Comment. Math. Helv. 77 (2002), 609–637.

[PR07] C. Pasnicu and M. Rørdam, *Purely infinite C^*-algebras of real rank zero*, J. Reine Angew. Math. 613 (2007), 51–73.

[Res06] G. Restorff, *Classification of Cuntz-Krieger algebras up to stable isomorphism*, J. Reine Angew. Math. 598 (2006), 185–210.

[RR07] G. Restorff and E. Ruiz, *On Rørdam's classification of certain C^*-algebras with one nontrivial ideal II*, Math. Scand. 101 (2007), 280–292.

[Ror97] M. Rørdam, *Classification of extensions of certain C^*-algebras by their six term exact sequences in K-theory*, Math. Ann. 308 (1997), 93–117.

[Sor13] A. Sørensen, *Geometric classification of simple graph algebras*, Ergodic Theory Dynam. Systems 33 (2013), no. 4, 1199–1220.

[TW07] A. Toms and W. Winter, *Strongly self-absorbing C*-algebras*, Trans. Amer. Math. Soc. 359 (2007), 3999–4029.

[Voi83] D. Voiculescu, *Asymptotically commuting finite rank unitary operators without commuting approximants*, Acta Sci. Math. (Szeged) 45 (1983), no. 1-4, 429–431.

Søren Eilers
Department of
Mathematical Sciences
University of Copenhagen
Universitetsparken 5
DK-2100 Copenhagen
Denmark

Xin Li
School of
Mathematical Sciences
Queen Mary University
of London
Mile End Road
London E1 4NS
United Kingdom

Efren Ruiz
Department of Mathematics
University of Hawaii, Hilo
200 W. Kawili St., Hilo
Hawaii, 96720-4091
USA

Some Endoscopic Properties of The Essentially Tame Jacquet-Langlands Correspondence

Kam-Fai Tam

Received: September 28. 2015

Communicated by Takeshi Saito

Abstract. Let F be a non-Archimedean local field of characteristic 0 and G be an inner form of the general linear group $G^* = \mathrm{GL}_n$ over F. We show that the rectifying character appearing in the essentially tame Jacquet-Langlands correspondence of Bushnell and Henniart for G and G^* can be factorized into a product of some special characters, called zeta-data in this paper, in the theory of endoscopy of Langlands and Shelstad. As a consequence, the essentially tame local Langlands correspondence for G can be described using admissible embeddings of L-tori.

2010 Mathematics Subject Classification: Primary 22E50; Secondary 11S37, 11F70.

Keywords and Phrases: essentially tame Jacquet-Langlands correspondence, inner forms, admissible pairs, zeta-data, endoscopy, admissible embeddings

Contents

1 Introduction

Let G^* be a general linear group over a non-Archimedean local field of characteristic 0, and G be an inner form of G^*. In this paper, we refine the results of the rectifying characters in the context of the essentially tame Jacquet-Langlands correspondence [BH11] by proving that each rectifying character admits a factorization into a product of characters called ζ-data, defined similarly to χ-data in [LS87], which are significant in describing the essentially tame local Langlands correspondence for G.

We know from [BH11] that the rectifying characters are quadratic characters that measure the difference between two correspondences for essentially tame supercuspidal representations of G and G^*: the representation theoretic one by matching the maximal simple types of the two groups, and the functorial one from the Jacquet-Langlands correspondence. On the representation theoretic side, the maximal simple types of G can be constructed using certain characters of its elliptic maximal tori, while on the functorial side, the Langlands parameters for G can be functorially lifted from the parameters of the same collection of characters.

Combining these results with our result on rectifying characters, we show that the essentially tame local Langlands correspondence for G can be described completely by admissible embeddings, defined in [LS87], of the L-groups of elliptic maximal tori into the L-group of G, generalizing an analogous description proved by the author [Tam] in the split case (when $G = G^*$).

1.1 Background

Let F be a non-Archimedean local field, G^* be the group GL_n defined over F, and G be an inner form of G^* defined over F. The set of F-points $G(F)$ of G is therefore isomorphic to $\mathrm{GL}_m(D)$ as a group, where D is a central division algebra over F of dimension d^2 and $m = n/d$.

Let $\mathcal{A}^2_m(D)$ (resp. $\mathcal{A}^2_n(F)$) be the discrete series of $G(F)$ (resp. $G^*(F)$), i.e., the set of equivalence classes of irreducible admissible representations that are essentially square integrable mod-center. The Jacquet-Langlands correspondence asserts a bijection

$$JL : \mathcal{A}^2_n(F) \to \mathcal{A}^2_m(D)$$

determined by a character relation (see (3.5)) between a representation in $\mathcal{A}^2_n(F)$ and its image in $\mathcal{A}^2_m(D)$. The existence of this bijection is known, starting from the case $n = 2$ [JL70], when $G(F)$ is the multiplicative group of the quaternion algebra over F. For arbitrary n, when $G(F)$ is the multiplicative group of a division algebra, the existence is proved by [Rog83]. The general situations are treated by [DKV84] in the characteristic zero case and by [Bad02], [BHL10] in the positive characteristic case.

Bushnell and Henniart describe in [BH11, (2.1)] the image of JL when it is restricted to the subset $\mathcal{A}^0_n(F)$ of supercuspidal representations. The image is the subset of representations in $\mathcal{A}^2_m(D)$, each of whose *parametric degree* is equal to n. We do not need the full definition of the parametric degree of a representation, so we only refer to [BH11, Section 2.7] for details. We only need to know that

- the parametric degree of a representation in $\mathcal{A}^2_m(D)$ is a positive integer divisor of n,

- a representation in $\mathcal{A}^2_m(D)$ is supercuspidal if its parametric degree is n; the converse is true in the split case (when $G = G^*$) but not in general, and
- the parametric degree is preserved under JL.

Furthermore, we can describe the image of JL of *each* representation $\pi \in \mathcal{A}^0_n(F)$ when π is *essentially tame*, a notion we will explain in Section 3.2. More precisely, if we let $\mathcal{A}^{\mathrm{et}}_m(D)$ (resp. $\mathcal{A}^{\mathrm{et}}_n(F)$) be the set of essentially tame representations in $\mathcal{A}^2_m(D)$ (resp. $\mathcal{A}^2_n(F)$) whose parametric degree is n, then we can describe completely the *essentially tame Jacquet-Langlands correspondence*:

$$JL : \mathcal{A}^{\mathrm{et}}_n(F) \to \mathcal{A}^{\mathrm{et}}_m(D),$$

as in [BH11, Theorem A].
To explain the theorem and describe JL completely, we require the notion of admissible characters from [How77]. In Section 3.1, we define the set $P_n(F)$ of (equivalence classes of) admissible pairs $(E/F, \xi) \in P_n(F)$, where E/F is a tamely ramified extension of degree n and ξ is a character of $E^\times$ admissible over F. This set bijectively parametrizes both $\mathcal{A}^{\mathrm{et}}_n(F)$ and $\mathcal{A}^{\mathrm{et}}_m(D)$ explicitly [BH11], using the theory of simple types of $G(F)$ developed in [BF85], [Gra07], [Séc04], [Séc05a], [Séc05b], [SS08], [BSS12] which generalizes the corresponding theory in the split case [BK93], [BH96] and the division algebra case [Zin92], [Bro96].
If we denote by

$$_F\Pi : P_n(F) \to \mathcal{A}^{\mathrm{et}}_n(F),\ (E/F, \xi) \mapsto {}_F\Pi_\xi \tag{1.1}$$

and

$$_D\Pi : P_n(F) \to \mathcal{A}^{\mathrm{et}}_m(D),\ (E/F, \xi) \mapsto {}_D\Pi_\xi \tag{1.2}$$

the above bijections, then Bushnell and Henniart proved in [BH11] that the composition

$$\nu : P_n(F) \xrightarrow{{}_F\Pi} \mathcal{A}^{\mathrm{et}}_n(F) \xrightarrow{JL} \mathcal{A}^{\mathrm{et}}_m(D) \xrightarrow{{}_D\Pi^{-1}} P_n(F)$$

maps an admissible pair $(E/F, \xi) \in P_n(F)$ to another pair of the form $(E/F, \xi \cdot {}_D\nu_\xi)$, where ${}_D\nu_\xi$ is a tamely ramified character ${}_D\nu_\xi$ of $E^\times$ depending on ξ. We borrow the terminology from [BH10] and call the character ${}_D\nu_\xi$ the *rectifier* of ξ for the essentially tame Jacquet-Langlands correspondence.

1.2 Main results

The main result of this paper is to relate the rectifier ${}_D\nu_\xi$ with a special set of characters, called ζ-data in this paper, introduced in the theory of endoscopy of Langlands and Shelstad [LS87]. The significance of ζ-data will be explained in the next section, together with a brief summary of the previous results of the author [Tam].

To describe the main result, we first assume that char$(F) = 0$ (see Remark 1.3 about this assumption). Let T be the F-torus such that $T(F) = E^{\times}$. We regard T as a maximal torus embedded in G. In contrast to the split case, we have to carefully choose the embedding $T \to G$ relative to an hereditary $\mathfrak{o}_F$-order in $G(F)$ associated to ξ. This will be explained in Section 2.5. Given this embedding, let $\Phi = \Phi(G, T)$ be the root system, which is invariant under the action of the absolute Galois group Γ_F of F if we view Φ as a subset of the character group of T. For each root $\lambda \in \Phi$, we denote by E_λ the fixed field of the stabilizer of λ in Γ_F, so that E_λ is a field extension of a Γ_F-conjugate of E. We recall from [LS87, Corollary 2.5.B] that ζ-data is a set of characters

$$\{\zeta_\lambda\} = \{\zeta_\lambda\}_{\lambda \in \Gamma_F \backslash \Phi},$$

where each ζ_λ is a character of $E_\lambda^{\times}$ satisfying the conditions in *loc. cit.* (and will be recalled in Section 5.1). Here λ ranges over a suitable subset of roots in Φ, denoted by $\Gamma_F \backslash \Phi$ for the moment, representing the Γ_F-orbits of Φ and such that E_λ is a field extension of E (but not just its conjugate).

The following theorem restates the main result, Theorem 5.5, in a simpler way.

Theorem 1.1. *Given a character ξ of $E^{\times}$ admissible over F.*

(i) There exists a set of ζ-data $\{\zeta_{\lambda,\xi}\}_{\lambda \in \Gamma_F \backslash \Phi}$ such that

$$_D\nu_\xi = \prod_{\lambda \in \Gamma_F \backslash \Phi} \zeta_{\lambda,\xi}|_{E^{\times}}.$$

(ii) The values of each $\zeta_{\lambda,\xi}$ can be expressed in terms of certain invariants, called t-factors in this paper, of the corresponding component in the complete symmetric decomposition of the finite symplectic modules associated to ξ (see the notations and definitions in Sections 4.2 and 3.7).

□

We explain statement (ii) of the above theorem. The finite symplectic modules appear in the respective constructions of the extended maximal simple types inducing $_F\Pi_\xi$ and $_D\Pi_\xi$ in (1.1) and (1.2) (see [BH11, (2.5.4)], or the summary in Section 3.5). Each of these modules admits an orthogonal decomposition, called a complete symmetric decomposition in this paper (Proposition 4.4), whose components are parametrized by the same set $\Gamma_F \backslash \Phi$ parameterizing the factors in (i) of the Theorem. The t-factors are, roughly speaking, defined by the symplectic signs attached to these components.

When proving Theorem 1.1, we pick a choice of characters $\{\zeta_{\lambda,\xi}\}_{\lambda \in \Gamma_F \backslash \Phi}$, where each character $\zeta_{\lambda,\xi}$ has values in terms of the t-factors of the corresponding component, the one indexed by λ. We then show that these characters constitute a set of ζ-data. Moreover, using the multiplicativity of t-factors, the product of these ζ-data, when restricted to $E^{\times}$, is equal to the rectifier $_D\nu_\xi$, whose values

are given in the First and Second Comparison Theorems of [BH11]. Hence our result refines the one in *loc. cit.*.
While the finite symplectic modules and their decompositions are also studied by the author in the split case [Tam], there are extra conditions on the components of these modules in the general case. These conditions come from the extra ramifications of the related compact subgroups in constructing the extended maximal simple types. The degrees of these ramifications depend on the residue degree $f(E/F)$ and, with other conditions similar to those in the split case, determine whether each component is trivial or not. This new phenomenon will be fully studied in Section 4. In particular, when E/F is totally ramified and ξ is fixed, the finite symplectic modules are isomorphic to each other for all inner forms of G^*, a fact already known in [BH11, Proposition 5.6].

1.3 RELATION WITH THE PREVIOUS RESULTS

The significance of the factorization of ${}_D\nu_\xi$ in Theorem 1.1(i) comes from [Tam], which proves an analogous factorization of the rectifier ${}_F\mu_\xi$ for the essentially tame local Langlands correspondence [BH05a].
We first recall from *loc. cit.* that the rectifier ${}_F\mu_\xi$ measures the difference between the "naïve correspondence" and the essentially tame Langlands correspondence for G^*; more precisely, the Langlands parameter of ${}_F\Pi_\xi$ defined in (1.1) is the induced representation

$$\mathrm{Ind}_{\mathcal{W}_E}^{\mathcal{W}_F}(\xi \cdot {}_F\mu_\xi) \tag{1.3}$$

of the Weil group $\mathcal{W}_F$ of F, where $\xi \cdot {}_F\mu_\xi$ is regarded as a character of $\mathcal{W}_E$ by class field theory [Tat79]. In [Tam, Theorem 1.1], the author proved that the rectifier ${}_F\mu_\xi$ admits a factorization

$${}_F\mu_\xi = \prod_{\lambda\in\Gamma_F\backslash\Phi} \chi_{\lambda,\xi}|_{E^\times},$$

where $\{\chi_{\lambda,\xi}\}_{\lambda\in\Gamma_F\backslash\Phi}$ is a set of χ-data, consisting of characters of $E_\lambda^\times$ satisfying the conditions similar to those of ζ-data (see Section 5.1).
With a collection of χ-data, we follow [LS87, Section 2.6] to construct an admissible embedding

$$I_{\{\chi_{\lambda,\xi}\}} : {}^LT \to {}^LG$$

of the L-group ${}^LT$ of the maximal torus T into the L-group ${}^LG$ of G^*. (For convenience, we call ${}^LT$ an L-torus in this paper.) Let $\tilde{\xi} : \mathcal{W}_F \to {}^LT$ be an L-homomorphism whose class is the parameter of the character ξ of $E^\times = T(F)$ under the local Langlands correspondence of the torus T, i.e., the Artin reciprocity for $E^\times$ [Tat79]. In a previous result of the author [Tam, Corollary 1.2], the Langlands parameter (1.3) of ${}_F\Pi_\xi$ is isomorphic to the composition

$$I_{\{\chi_{\lambda,\xi}\}} \circ \tilde{\xi} : \mathcal{W}_F \to {}^LT \to {}^LG \xrightarrow{\text{natural proj.}} \mathrm{GL}_n(\mathbb{C})$$

as a representation of $\mathcal{W}_F$. In other words, the essentially tame local Langlands correspondence for G^* can be described by admissible embeddings of L-tori.
A set of ζ-data is the 'difference' of two sets of χ-data, in the sense that, given a set $\{\chi_\lambda\}$ of χ-data, we have

$$\{\chi'_\lambda\} \text{ is another set of } \chi\text{-data} \quad \Leftrightarrow \quad \{\chi_\lambda(\chi'_\lambda)^{-1}\} \text{ is a set of } \zeta\text{-data.}$$

If we define the local Langlands correspondence for G as the composition of the local Langlands correspondence for G^* and the Jacquet-Langlands correspondence JL, then we can express the Langlands parameter of ${}_D\Pi_\xi$ using an admissible embedding of L-tori, as follows.

Corollary 1.2. *Let $\{\chi_{\lambda,\xi}\}$ and $\{\zeta_{\lambda,\xi}\}$ be respectively the χ-data and the ζ-data associated to an admissible character ξ. The Langlands parameter*

$$\mathrm{Ind}_{\mathcal{W}_E}^{\mathcal{W}_F}(\xi \cdot {}_F\mu_\xi \cdot {}_D\nu_\xi)$$

of ${}_D\Pi_\xi$ is isomorphic to

$$I_{\{\chi_{\lambda,\xi}\cdot\zeta_{\lambda,\xi}\}} \circ \tilde{\xi} : \mathcal{W}_F \to {}^LT \to {}^LG \xrightarrow{\text{natural proj.}} \mathrm{GL}_n(\mathbb{C})$$

as a representation of $\mathcal{W}_F$. □

Hence analogously we can describe the essentially tame local Langlands correspondence for G by admissible embeddings of L-tori.
As a consequence, we show in Proposition 5.7 that the factorization of ${}_D\nu_\xi$ in Theorem 1.1(i) is functorial, in the following sense. Let K/F be an intermediate extension of E/F, so that if the pair $(E/F,\xi)$ is admissible over F, then $(E/K,\xi)$ is admissible over K by definition. We denote the centralizer of $K^\times$ in $G(F)$ by $GL_{m_K}(D_K)$, where D_K is a K-division algebra and m_K a positive integer. If $\{\zeta_{\lambda,\xi}\}$ is the set of ζ-data associated to ξ, then the partial product

$$\prod_{\lambda\in\Gamma_F\backslash\Phi,\ \lambda|_{K^\times}\neq 1} \zeta_{\lambda,\xi}|_{E^\times}.$$

(a product similar to Theorem 1.1.(i), with factors ranging over the characters being non-trivial on $K^\times$) is the rectifier ${}_{D_K}\nu_\xi$ of ξ over K.

Remark 1.3. We would like to remark on the condition of the characteristic $\mathrm{char}(F) = 0$, as we also did in [Tam, Remark 1.3]. The readers should be aware that the works of [JL70], [DKV84], [Bad02], [BH11] make the Jacquet-Langlands correspondence valid for local fields of arbitrary characteristic. In our paper, we apply the condition $\mathrm{char}(F) = 0$ only because we refer to the theory of endoscopy from [LS87], [KS99]. However, we do not actually need this condition for the part of the theory that we allude to, which is about the admissible embeddings of L-tori. In [Tam, Section 6] (or rather [LS87, Section 2.5]), we see that these kind of embeddings can be defined without any condition of $\mathrm{char}(F)$. Therefore, the condition $\mathrm{char}(F) = 0$ in this paper should be treated as a mild condition. □

1.4 Notations

Throughout the paper, F denotes a non-Archimedean local field of characteristic 0. Its ring of integers is $\mathfrak{o}_F$ with the maximal ideal $\mathfrak{p}_F$. The residue field $\mathbf{k}_F = \mathfrak{o}_F/\mathfrak{p}_F$ has q elements and is of characteristic p. We denote by $v_F : F^\times \to \mathbb{Z}$ the discrete valuation on F. We denote by Γ_F the absolute Galois group of F, and by $\mathcal{W}_F$ the Weil group of F.
The multiplicative group $F^\times$ decomposes into a product of subgroups

$$\langle \varpi_F \rangle \times \boldsymbol{\mu}_F \times U_F^1.$$

They are namely the group generated by a prime element ϖ_F, the group $\boldsymbol{\mu}_F$ of roots of unity of order prime to p, and the 1-unit group $U_F^1 := 1 + \mathfrak{p}_F$. We will identify $\boldsymbol{\mu}_F$ with $\mathbf{k}_F^\times$ in the natural way. We then write $U_F = U_F^0 := \boldsymbol{\mu}_F \times U_F^1$ and $U_F^i := 1 + \mathfrak{p}_F^i$ for each positive integer i. Let $\boldsymbol{\mu}_n$ be the group of nth roots of unity in the algebraic closure $\bar{F}$ of F, and z_n be a choice of primitive nth root in $\boldsymbol{\mu}_n$.
The F-level of a character ξ of $F^\times$ is the smallest integer $a \geq -1$ such that $\xi|_{U_F^{a+1}}$ is trivial. A character ξ of $F^\times$ is called unramified if $\xi|_{U_F}$ is trivial, or equivalently, if its F-level is -1. It is called tamely ramified if $\xi|_{U_F^1}$ is trivial, or equivalently, if its F-level is 0.
Given a field extension E/F, we denote its ramification index by $e = e(E/F)$ and its residue degree by $f = f(E/F)$. We also denote by $tr_{E/F}$ and norm $N_{E/F}$ the trace and norm respectively.
We fix an additive character ψ_F of F of level 0, which means that ψ_F is trivial on $\mathfrak{p}_F$ but is non-trivial on $\mathfrak{o}_F$. Hence $\psi_F|_{\mathfrak{o}_F}$ induces a non-trivial character of $\mathbf{k}_F$. We write $\psi_E = \psi_F \circ \mathrm{tr}_{E/F}$.
Suppose that A is a central simple algebra over F. We denote the reduced trace by $\mathrm{trd}_{A/F}$ and the reduced norm by $\mathrm{Nrd}_{A/F}$.
Given a set X, we denote its cardinality by $\#X$. If H is a group and X is a H-set, then we denote the action of $h \in H$ on $x \in X$ by $x \mapsto {}^h x$. The set of H-orbits is denoted by $H\backslash X$. If π is a representation of H (over a given field), we denote its equivalence class by (H, π).

2 Some basic setups

2.1 Root system

Given a field extension E/F of degree n, we let T be the induced torus $\mathrm{Res}_{E/F}\mathbb{G}_m$. We embed T into G as an elliptic maximal torus, and denote the image still by T. The choice of this embedding will be specific in Section 2.5, but at this moment this choice is irrelevant. Let $\Phi = \Phi(G,T)$ be the root system of T in G. Following [Tam, Section 3.1], we can denote each root in Φ by $\left[\begin{smallmatrix} g \\ h \end{smallmatrix}\right]$ where $g = g\Gamma_E$ and $h = h\Gamma_E$ are distinct cosets in Γ_F/Γ_E. (We use the same notation g for an element in Γ_F and its Γ_E-cosets, for notation convenience.) The Γ_F-action on Φ is given by $x \cdot \left[\begin{smallmatrix} g \\ h \end{smallmatrix}\right] = \left[\begin{smallmatrix} xg \\ xh \end{smallmatrix}\right]$. For each root

$\lambda \in \Phi$, we denote by $[\lambda]$ its Γ_F-orbit $\mathcal{W}_F\lambda$. Each Γ_F-orbit contains a root of the form $\left[\begin{smallmatrix}1\\g\end{smallmatrix}\right]$ for some non-trivial coset $g \in \Gamma_F/\Gamma_E$.
For each root $\lambda \in \Phi$, we denote the stabilizers $\{g \in \Gamma_F | {}^g\lambda = \lambda\}$ and $\{g \in \Gamma_F | {}^g\lambda = \pm\lambda\}$ by Γ_λ and $\Gamma_{\pm\lambda}$ respectively and their fixed fields by E_λ and $E_{\pm\lambda}$ respectively. We call a root λ *symmetric* if $[E_\lambda : E_{\pm\lambda}] = 2$, and *asymmetric* otherwise. Equivalently, λ is symmetric if and only if λ and $-\lambda$ are in the same Γ_F-orbit. Note that the symmetry of Φ is preserved by the Γ_F-action. Let

(i) $\Gamma_F\backslash\Phi_{\mathrm{sym}}$ be the set of Γ_F-orbits of symmetric roots,

(ii) $\Gamma_F\backslash\Phi_{\mathrm{asym}}$ be the set of Γ_F-orbits of asymmetric roots, and

(iii) $\Gamma_F\backslash\Phi_{\mathrm{asym}/\pm}$ be the set of equivalence classes of asymmetric Γ_F-orbits by identifying $[\lambda]$ and $[-\lambda]$.

We denote by $(\Gamma_E\backslash\Gamma_F/\Gamma_E)'$ the collection of non-trivial double cosets, and by $[g]$ the double coset $\Gamma_E g \Gamma_E$. We can deduce the following proposition easily.

Proposition 2.1. *The map*

$$\Gamma_F\backslash\Phi \to (\Gamma_E\backslash\Gamma_F/\Gamma_E)',\ [\lambda] = \mathcal{W}_F \left[\begin{smallmatrix}1\\g\end{smallmatrix}\right] \mapsto [g],$$

is a bijection between the set $\Gamma_F\backslash\Phi$ of Γ_F-orbits of the root system Φ and the set $(\Gamma_E\backslash\Gamma_F/\Gamma_E)'$ of non-trivial double cosets.

We can therefore call $g \in \Gamma_F$ *symmetric* if $[g] = [g^{-1}]$, and *asymmetric* otherwise, so that the bijection in Proposition 2.1 preserves symmetries on both sides. Let

(i) $(\Gamma_E\backslash\Gamma_F/\Gamma_E)_{\mathrm{sym}}$ be the set of symmetric non-trivial double cosets,

(ii) $(\Gamma_E\backslash\Gamma_F/\Gamma_E)_{\mathrm{asym}}$ be the set of asymmetric non-trivial double cosets, and

(iii) $(\Gamma_E\backslash\Gamma_F/\Gamma_E)_{\mathrm{asym}/\pm}$ be the set of equivalence classes of $(\Gamma_E\backslash\Gamma_F/\Gamma_E)_{\mathrm{asym}}$ by identifying $[g]$ with $[g^{-1}]$.

We choose subsets $\mathcal{D}_{\mathrm{sym}}$ and $\mathcal{D}_{\mathrm{asym}/\pm}$ of representatives in Γ_F/Γ_E of $(\Gamma_E\backslash\Gamma_F/\Gamma_E)_{\mathrm{sym}}$ and $(\Gamma_E\backslash\Gamma_F/\Gamma_E)_{\mathrm{asym}/\pm}$ respectively, and write

$$\mathcal{D}_{\mathrm{asym}} = \mathcal{D}_{\mathrm{asym}/\pm} \sqcup \{g^{-1} | g \in \mathcal{D}_{\mathrm{asym}/\pm}\}.$$

We also choose subsets $\mathcal{R}_{\mathrm{sym}}$ and $\mathcal{R}_{\mathrm{asym}/\pm}$ of representatives in Φ of orbits in $\Gamma_F\backslash\Phi_{\mathrm{sym}}$ and $\Gamma_F\backslash\Phi_{\mathrm{asym}/\pm}$ respectively such that every root $\lambda \in \mathcal{R}_\pm := \mathcal{R}_{\mathrm{sym}} \sqcup \mathcal{R}_{\mathrm{asym}/\pm}$ is of the form $\left[\begin{smallmatrix}1\\g\end{smallmatrix}\right]$ for some $g \in \mathcal{D}_\pm := \mathcal{D}_{\mathrm{sym}} \sqcup \mathcal{D}_{\mathrm{asym}/\pm}$, and write

$$\mathcal{R}_{\mathrm{asym}} = \mathcal{R}_{\mathrm{asym}/\pm} \sqcup (-\mathcal{R}_{\mathrm{asym}/\pm}).$$

Hence $\mathcal{R}_{\mathrm{sym}}$, $\mathcal{R}_{\mathrm{asym}}$, and $\mathcal{R}_{\mathrm{asym}/\pm}$ correspond bijectively to $\mathcal{D}_{\mathrm{sym}}, \mathcal{D}_{\mathrm{asym}}$, and $\mathcal{D}_{\mathrm{asym}/\pm}$ respectively by the identification in Proposition 2.1. Denote $E_g := E_\lambda$ and $E_{\pm g} := E_{\pm\lambda}$. Notice that $E_g = E({}^gE)$, composite field of E and ${}^gE$.

2.2 Galois groups

Let E/F be a field extension of degree n. In most of the paper, we assume that E/F is tamely ramified, which means that p is coprime to e. By [Lan94, II.§5], we can choose ϖ_E and ϖ_F such that

$$\varpi_E^e = z_{E/F}\varpi_F, \text{ for some } z_{E/F} \in \boldsymbol{\mu}_E. \tag{2.1}$$

Choose in $\bar{F}^\times$ a primitive eth root of unity z_e and an eth root $z_{E/F,e}$ of $z_{E/F}$. (We do not require that $z_{E/F,e}^a = z_e$, if a is the multiplicative order of $z_{E/F}$.) Denote $L = E[z_e, z_{E/F,e}]$ and $l = [L : E]$. With the choices of ϖ_F and ϖ_E as in (2.1), we define the following F-automorphisms on L.

(i) $\phi : z \mapsto z^q$, for all $z \in \boldsymbol{\mu}_L$, and $\phi : \varpi_E \mapsto z_\phi \varpi_E$.

(ii) $\sigma : z \mapsto z$, for all $z \in \boldsymbol{\mu}_L$, and $\sigma : \varpi_E \mapsto z_e \varpi_E$.

Here z_ϕ lies in $\boldsymbol{\mu}_E$ satisfying $(z_\phi\varpi_E)^e = z_{E/F}^q\varpi_F$. More generally, we write ${}^{\phi^i}\varpi_E = z_{\phi^i}\varpi_E$ where $z_{\phi^i} = z_\phi^{1+q+\cdots+q^{i-1}}$ is an eth root of $z_{E/F}^{q^i-1}$.
Therefore, $\Gamma_{L/F} = \langle\sigma\rangle \rtimes \langle\phi\rangle$ with relation $\phi\circ\sigma\circ\phi^{-1} = \sigma^q$. Suppose that $\Gamma_{L/E} = \langle\sigma^c\phi^f\rangle$ for some integer c satisfying the condition:

$$e \text{ divides } c\left(\frac{q^{fl}-1}{q^f-1}\right).$$

We can choose

$$\{\sigma^i\phi^j | i = 0,\dots,e-1,\ j = 0,\dots,f-1\}$$

as coset representatives for the quotient $\Gamma_{E/F} = \Gamma_F/\Gamma_E$. Moreover, elements in a fixed double coset are of the form $[\sigma^i\phi^j]$ with a fixed j mod f.

Proposition 2.2 ([Tam, Proposition 3.3]). *The double coset $[g] = [\sigma^i\phi^j]$ is symmetric only if $j = 0$ or, when f is even, $j = f/2$.* □

We call those symmetric $[\sigma^i]$ ramified and those symmetric $[\sigma^i\phi^{f/2}]$ unramified, and denote by $(\Gamma_E\backslash\Gamma_F/\Gamma_E)_{\mathrm{sym-ram}}$ and $(\Gamma_E\backslash\Gamma_F/\Gamma_E)_{\mathrm{sym-unram}}$ respectively the collections of symmetric ramified and symmetric unramified double cosets. We provide several useful results concerning the parity of certain subsets in $\Gamma_E\backslash\Gamma_F/\Gamma_E$.

Proposition 2.3 ([Tam, Propositions 3.4 and 3.5]). *(i) If $[g]$ is symmetric unramified, then the degree $[E_g : E]$ is odd.*

(ii) The parity of $\#(\Gamma_E\backslash\Gamma_F/\Gamma_E)_{\mathrm{sym-unram}}$ is equal to that of $e(f-1)$. □

Lemma 2.4. *Suppose that f is even. The following are equivalent.*

(i) There exists $\sigma^i\phi^{f/2} \in W_{F[\varpi_E]}$ for some i.

(ii) $z_{\phi^{f/2}}$ *is an* e*th root of unity.*

(iii) $z_{E/F} \in K_+$, *where* K_+/F *is unramified of degree* $f/2$.

(iv) $f_\varpi := [E : F[\varpi_E]]$ *is even.*

Proof. (i) is equivalent to (ii) since ${}^{\sigma^i\phi^{f/2}}\varpi_E = z_e^i z_{\phi^{f/2}}\varpi_E$. To show that (iii) implies (ii), we recall that $z_{\phi^{f/2}}$ is an eth root of $z_{E/F}^{q^{f/2}-1}$. If $z_{E/F} \in K_+$, then $z_{E/F}^{q^{f/2}-1} = 1$ and $z_{\phi^{f/2}}$ is an eth root of unity. The converse is similar. To show the equivalence of (iii) and (iv), we notice that $f(F[\varpi_E]/F) = f(F[z_{E/F}]/F) = f/f_\varpi$. Hence that $F[z_{E/F}] \subseteq K_+$ is equivalent to that f_ϖ is even. $\square$

Lemma 2.5. *Suppose that* $g = \sigma^i\phi^{f/2}$ *satisfies the conditions in Lemma 2.4.*

(i) The double coset $[\sigma^i\phi^{f/2}]$ *is automatically symmetric.*

(ii) The set $(\Gamma_E\backslash\Gamma_{F[\varpi_E]}/\Gamma_E)_{\mathrm{sym-unram}} = (\Gamma_E\backslash\Gamma_F/\Gamma_E)_{\mathrm{sym-unram}} \cap (\Gamma_E\backslash\Gamma_{F[\varpi_E]}/\Gamma_E)$ *contains a single element* $[\sigma^i\phi^{f/2}]$.

Proof. For (i), we consider the actions of $\sigma^i\phi^{f/2}$ and its inverse $(\sigma^i\phi^{f/2})^{-1}$ on E. We certainly have ${}^{\sigma^i\phi^{f/2}}\varpi_E = {}^{(\sigma^i\phi^{f/2})^{-1}}\varpi_E = \varpi_E$ by definition. We also have ${}^{\sigma^i\phi^{f/2}}z = {}^{(\sigma^i\phi^{f/2})^{-1}}z = z^{q^{f/2}}$ for all $z \in \boldsymbol{\mu}_E$. Therefore $\sigma^i\phi^{f/2}\Gamma_E = (\sigma^i\phi^{f/2})^{-1}\Gamma_E$ and in particular $[\sigma^i\phi^{f/2}] = [(\sigma^i\phi^{f/2})^{-1}]$. For (ii), we know by Lemma 2.4.(ii) that the double coset is the one containing $\sigma^i\phi^{f/2}$ where $z_{\phi^{f/2}} = z_e^{-i}$. $\square$

Proposition 2.6. *When* f *is even, the parity of the cardinality of* $(\Gamma_E\backslash\Gamma_F/\Gamma_E)_{\mathrm{sym-unram}} - \Gamma_{F[\varpi_E]}$ *is equal to* $e + f_\varpi - 1$.

Proof. Recall

(i) by Lemma 2.4 that there exists $\sigma^i\phi^{f/2} \in W_{F[\varpi_E]}$ if and only if f_ϖ is even, and

(ii) by Proposition 2.3 that the parity of the number of symmetric $[\sigma^i\phi^{f/2}]$ is the same as that of e.

By combining these facts, we have the assertion. $\square$

2.3 Division algebra

Let D be a division algebra over F of dimension n^2. Denote its unique maximal order by $\mathfrak{o}_D$ and the maximal ideal of $\mathfrak{o}_D$ by $\mathfrak{p}_D$. Suppose that the Hasse-invariant of D is $h = h(D)$, so that $\gcd(n, h) = 1$. By [Rei03, (14.5) Theorem], we can choose a primitive $(q^n - 1)$th root z of unity in D and a uniformizer ϖ_D such that

$$\varpi_D^n = \varpi_F \text{ and } \varpi_D z \varpi_D^{-1} = z^{q^h}. \tag{2.2}$$

We write $K_n = F[z]$ and $\boldsymbol{\mu}_D = \langle z \rangle$, then K_n is a maximal unramified extension (of degree n) in D, and $\boldsymbol{\mu}_D$ is a group of roots of unity of order $q^n - 1$, both defined up to conjugacy by $D^\times$. Therefore, the conjugation of ϖ_D acts on K_n as the hth power of the Frobenius automorphism, i.e.,

$$\varpi_D u \varpi_D^{-1} = {}^{\phi^h}u \text{ for all } u \in K_n.$$

We write $U_D^i := 1 + \mathfrak{p}_D^i$ for all positive integer i. The multiplicative subgroup $D^\times$ hence decomposes into a semi-direct product

$$(\langle \varpi_D \rangle \ltimes \boldsymbol{\mu}_D) \ltimes U_D^1.$$

Let E/F be a tamely ramified field extension of degree n, and let K be the maximal unramified sub-extension in E/F. We assume that $K \subseteq K_n$ and that the uniformizers ϖ_E and ϖ_F satisfy $\varpi_E^e = z_{E/F}\varpi_F$ as in (2.1) for some $z_{E/F} \in \boldsymbol{\mu}_E$. If we define $z_{D/E} \in \boldsymbol{\mu}_D = \boldsymbol{\mu}_{K_n}$ to be a solution of

$$N_{K_n/K}(z_{D/E}) = z_{E/F}, \tag{2.3}$$

then we may take $\varpi_E = \varpi_D^f z_{D/E}$ and this defines an embedding of E into D over F. Note that from (2.2)

$$z\varpi_D^i z^{-1} = z^{1-q^{hi}}\varpi_D^i \tag{2.4}$$

for all $z \in \boldsymbol{\mu}_E = \boldsymbol{\mu}_K$ and all $i \in \mathbb{Z}$.

2.4 Hereditary orders in central simple algebra

If G is an F-inner form of $G^* = \mathrm{GL}_n$, then $G(F) = A^\times$, where A be a central simple algebra over F. By Wedderburn Theorem [Rei03, (7.4)Theorem], A is isomorphic to $\mathrm{Mat}_m(D)$, where D is a division algebra of F-dimension d^2 and $md = n$. Therefore, $G(F) \cong \mathrm{GL}_m(D)$. Any field extension of degree n can be embedded into A as a maximal subfield in A, and any two such embeddings are conjugate under $G(F)$.

Let $\mathfrak{A}$ be an $\mathfrak{o}_F$-hereditary order in A, $\mathfrak{P}_\mathfrak{A}$ be its Jacobson radical, and $\mathfrak{K}_\mathfrak{A}$ be the $G(F)$-normalizer of $\mathfrak{A}^\times$. If $\mathfrak{A}$ is principal, in the sense that there exists $\varpi_\mathfrak{A} \in \mathfrak{K}_\mathfrak{A}$ such that $\varpi_\mathfrak{A}\mathfrak{A} = \mathfrak{A}\varpi_\mathfrak{A} = \mathfrak{P}_\mathfrak{A}$, then the valuation $v_\mathfrak{A} : \mathfrak{K}_\mathfrak{A} \to \mathbb{Z}$ is defined by $x\mathfrak{A} = \mathfrak{A}x = \mathfrak{P}_\mathfrak{A}^{v_\mathfrak{A}(x)}$ for all $x \in \mathfrak{K}_\mathfrak{A}$. We also write $U_\mathfrak{A} = U_\mathfrak{A}^0 = \mathfrak{A}^\times$, $U_\mathfrak{A}^i = 1 + \mathfrak{P}_\mathfrak{A}^i$ for each positive integer i, $U_\mathfrak{A}^x = U_\mathfrak{A}^{\lceil x \rceil}$ for all $x \in \mathbb{R}_{\geq 0}$, and

$$U_\mathfrak{A}^{x+} = \bigcup_{x \in \mathbb{R}_{\geq 0},\, y > x} U_\mathfrak{A}^y.$$

Suppose that E_0 is a subfield in A and $\mathfrak{A}$ is E_0-pure, i.e., ${E_0}^\times \subseteq \mathfrak{K}_\mathfrak{A}$, then we define the ramification index $e(\mathfrak{A}/\mathfrak{o}_{E_0})$ to be the integer e satisfying $v_\mathfrak{A}|_{{E_0}^\times} = ev_{E_0}$. We therefore have

$$\mathfrak{p}_{E_0}^i \mathfrak{P}_\mathfrak{A}^j = \mathfrak{P}_\mathfrak{A}^{ie+j}$$

and

$$\mathfrak{p}_{E_0}^i \cap \mathfrak{P}_{\mathfrak{A}}^j = \mathfrak{p}_{E_0}^{\max\{i,j/e\}}$$

for all $i, j \in \mathbb{Z}$.

In the split case $G = G^*$, i.e., when $A = A^* = \mathrm{Mat}_n(F)$, we denote the hereditary order $\mathfrak{A}$ by $\mathfrak{A}^*$.

2.5 Embedding conditions

Suppose we fix an F-embedding $E_0 \hookrightarrow A$, let E_0 and $\mathfrak{A}$ be as in the previous section, and write A_0 the centralizer of E_0 in A. We can restrict the embedding to $E_0^\times \hookrightarrow G(F)$, and denote the centralizer by $A_0^\times = Z_{G(F)}(E_0^\times)$. Under the above setup, there are many choices of $\mathfrak{A}$ among its $Z_{G(F)}(E_0^\times)$-conjugacy class. In this paper, we assume the conditions (i)-(iii) below, all adopted from [BH11], to fix a unique $\mathfrak{A}$.

(i) [BH11, Section 3.2] If E/E_0 is an unramified extension in A_0 such that $[E : F] = n$, then we require that $\mathfrak{A}$ is E-pure, i.e., $E^\times \subseteq \mathfrak{K}_{\mathfrak{A}}$.

Let $\mathfrak{A}_0$ be the centralizer of E_0 in $\mathfrak{A}$, i.e., $\mathfrak{A}_0 = \mathfrak{A} \cap A_0$, which is a hereditary $\mathfrak{o}_{E_0}$-order in A_0, with Jacobson radical $\mathfrak{P}_{\mathfrak{A}_0} = \mathfrak{P}_{\mathfrak{A}} \cap A_0$.

(ii) [BH11, (2.3.1)(2)] There exists a fixed integer $e(\mathfrak{A}/\mathfrak{A}_0) \geq 1$ such that

$$\mathfrak{P}_{\mathfrak{A}}^k \cap A_0 = \mathfrak{P}_{\mathfrak{A}_0}^{k/e(\mathfrak{A}/\mathfrak{A}_0)} \text{ for every } k \in \mathbb{Z}. \tag{2.5}$$

We say that $\mathfrak{A}$ is the *canonical continuation* of $\mathfrak{A}_0$ in A. Under (ii), we have moreover $\mathfrak{K}_{\mathfrak{A}} \cap A_0 = \mathfrak{K}_{\mathfrak{A}_0}$.

(iii) [BH11, (2.3.2)] $\mathfrak{A}_0$ is a maximal hereditary $\mathfrak{o}_{E_0}$-order in A_0, i.e., $e(\mathfrak{A}_0/\mathfrak{o}_{E_0}) = 1$.

Under these conditions, $\mathfrak{A}$ is the unique E_0-pure hereditary order in A such that $\mathfrak{A} \cap A_0 = \mathfrak{A}_0$. Moreover, $\mathfrak{A}$ is maximal among all E_0-pure hereditary orders in A, and both $\mathfrak{A}$ and $\mathfrak{A}_0$ are principal (by [BH11, the remark after (2.3.2)]).

By [Zin99, 0. Theorem], if the $\mathfrak{o}_D$-period of $\mathfrak{A}$ is denoted by $r = r(\mathfrak{A}) = e(\mathfrak{A}/\mathfrak{o}_D)$, i.e. $\varpi_D \mathfrak{A} = \mathfrak{P}_{\mathfrak{A}}^r$, then we have an isomorphism

$$\mathfrak{A}/\mathfrak{P}_{\mathfrak{A}} \cong \mathrm{Mat}_s(\mathbf{k}_D)^r, \tag{2.6}$$

where $s = s(\mathfrak{A}) = f/e(\mathfrak{A}/\mathfrak{o}_E)$. Once A (and hence D) is fixed, the integers r and s depend only on E; indeed $s = s(E/F) = \gcd(f, m)$ and $r = r(E/F) = e/\gcd(d, e) = m/s$.

If K is an intermediate subfield in E/F, we write $f_K = f(E/K)$ and $e_K = e(E/K)$. By [Zin99, 1. Prop.] the centralizer $A_K = Z_A(K)$ is isomorphic to $\mathrm{Mat}_{m_k}(D_K)$, where D_K is a division algebra over K of degree d_K^2, with

$$d_K = \frac{d}{\gcd(d, n(K/F))} \text{ and } m_K = \gcd(m, n(E/K)).$$

Let $\mathfrak{A}_K$ be the centralizer of K in $\mathfrak{A}$. It is routine to check that $(E_0K, \mathfrak{A}_K)$ is a canonical continuation of $\mathfrak{A}_{E_0K}$ and is also E-pure, i.e., (i)-(iii) are satisfied when we change our base field from F to K. We therefore have an isomorphism

$$\mathfrak{A}_K/\mathfrak{P}_{\mathfrak{A}_K} \cong \mathrm{Mat}_{s_K}(\mathbf{k}_{D_K})^{r_K}, \tag{2.7}$$

where $r_K = e_K/\gcd(d_K, e_K)$ and $s_K = \gcd(f_K, m_K)$.

2.6 Characters

Let $\mathrm{Nrd}_{A/F} : A^\times \to F^\times$ be the reduced norm of $G(F) = A^\times$. By [Rei03, (33.4)Theorem] and [NM43, Satz 2], $\mathrm{Nrd}_{A/F}$ is surjective and its kernel is the commutator subgroup of $G(F)$. Therefore, any character of $G(F)$ factors through $\mathrm{Nrd}_{A/F}$.

We define the $\mathfrak{A}$-level of a character ξ of $A^\times$ as the smallest integer $a \geq -1$ such that $\xi|_{U_{\mathfrak{A}}^{a+1}}$ is trivial. This is analogous to the F-level of a character of $F^\times$ defined in Section 1.4.

Proposition 2.7. *If the F-level of a character ξ of $F^\times$ is r, then the $\mathfrak{A}$-level of the character $\xi \circ \mathrm{Nrd}_{A/F}$ is $r \cdot e(\mathfrak{A}/\mathfrak{o}_F)$.*

Proof. It suffices to show that

$$\mathrm{Nrd}_{A/F}(U_{\mathfrak{A}}^{r\cdot e(\mathfrak{A}/\mathfrak{o}_F)}) = U_F^r \text{ for all } r \in \mathbb{R}_{\geq 0}.$$

This follows from [BF85, (2.8.3)]. □

3 The essentially tame Jacquet-Langlands correspondence

3.1 Admissible characters

We recall the definition of admissible characters in [How77], [Moy86]. Given a tamely ramified finite extension E/F, let ξ be a character of $E^\times$. We call a pair $(E/F, \xi)$ *admissible* over F if ξ is an admissible character over F, i.e., for every intermediate subfield K between E/F,

(i) if ξ factors through the norm $N_{E/K}$, then $E = K$;

(ii) if $\xi|_{U_E^1}$ factors through $N_{E/K}$ then E/K is unramified.

Two admissible pairs $(E/F, \xi)$ and $(E'/F, \xi')$ are called F-equivalent if there is $g \in \Gamma_F$ such that ${}^gE = E'$ and ${}^g\xi = \xi'$. Let $P_n(F)$ be the set of F-equivalence classes of the admissible pair $(E/F, \xi)$, with each class in $P_n(F)$ still denoted by $(E/F, \xi)$ for convenience.

Every admissible character ξ admits a factorization (see [How77, Corollary of Lemma 11] or [Moy86, Lemma 2.2.4])

$$\xi = \xi_{-1}(\xi_0 \circ N_{E/E_0}) \cdots (\xi_t \circ N_{E/E_t})(\xi_{t+1} \circ N_{E/F}), \tag{3.1}$$

where, in the notations above, the decreasing sequence of fields

$$E = E_{-1} \supseteq E_0 \supsetneq E_1 \supsetneq \cdots \supsetneq E_t \supsetneq E_{t+1} = F.$$

and the increasing E-levels $a_{-1} = 0 < a_0 < a_1 < \cdots < a_t \leq a_{t+1}$ of the characters $\xi_k \circ N_{E/E_k}$, $k = 0, \dots, t+1$, are uniquely determined. We call the E-levels a_k the *jumps* of ξ and call the collection $\{E_k, a_k | k = 0, \dots, t\}$ the *jump data* of ξ. By convention, when $E_0 = E$, we replace $(\xi_0 \circ N_{E/E_0})\xi_{-1}$ by ξ_0 and assume that ξ_{-1} is trivial; otherwise, ξ_{-1} is tamely ramified and E/E_0 is unramified [Moy86, Defnition 2.2.3]. There are certain generic conditions imposed on the jump data of the character by its admissibility, but we do not need them fully in this paper. We refer the interested reader to [Moy86, Definition 2.2.3] and [BH10, Section 8.2] for these conditions (and when E/F is totally ramified, see also [BH05b, Section 1]), and only use one of their consequences in (3.11).
We fix a (non-canonical) choice of ξ_{-1} in the factorization (3.1) as follows. We fix a choice of the wild component ξ_w of ξ to be the product

$$(\xi_0 \circ N_{E/E_0}) \cdots (\xi_t \circ N_{E/E_t})(\xi_{t+1} \circ N_{E/F})$$

which satisfies

$$\xi_w(\varpi_F) = 1 \text{ and } \xi_w \text{ has a } p\text{-power order} \tag{3.2}$$

(see [BH11, Lemma 1 of Section 4.3]), and define the tame component of ξ to be $\xi_{-1} = \xi\xi_w^{-1}$. We write

$$\Xi = \Xi(\xi) = \xi_0(\xi_1 \circ N_{E_0/E_1}) \cdots (\xi_{t+1} \circ N_{E_0/F}), \tag{3.3}$$

such that $\xi_w = \Xi \circ N_{E/E_0}$.
Suppose that E_0/F is a tamely ramified extension of degree dividing n and E/E_0 is unramified. Let Ξ be a character of $U^1_{E_0}$. Following [BH05a, Section 1.3], we call $(E_0/F, \Xi)$ an *admissible 1-pair* over F if Ξ does not factor through any norm $N_{E_0/K}$ with $F \subseteq K \subsetneq E_0$. We denote by $P^1_n(F)$ the set of F-equivalence classes of these pairs. Therefore, the map

$$P_n(F) \to P^1_n(F),\ (E/F, \xi) \mapsto (E_0/F, \Xi(\xi)|_{U^1_{E_0}}),$$

is well-defined and surjective. Notice that we can define the jump-data of a 1-pair, such that the jump-data of an admissible pair is the same as that of its associated 1-pair.

3.2 The correspondences

Let G^* be GL_n defined over F, and G be an inner form of G^* whose F-point is isomorphic to $\mathrm{GL}_m(D)$ for some central F-division algebra D of dimension d^2 and $n = md$. Let $\mathcal{A}^2_n(F)$ (resp. $\mathcal{A}^2_m(D)$) be the collection of equivalence classes of irreducible smooth representations of $G^*(F)$ (resp. $G(F)$) which are essentially square-integrable modulo center. The Jacquet-Langlands correspondence is a canonical bijection

$$JL : \mathcal{A}^2_n(F) \to \mathcal{A}^2_m(D) \tag{3.4}$$

between the two collections determined by a character relation between $\pi \in \mathcal{A}^2_n(F)$ and its image $JL(\pi) \in \mathcal{A}^2_m(D)$: for every pair of semi-simple elliptic regular elements (g, g^*), where $g \in G(F)$ and $g^* \in G^*(F)$, with the same reduced characteristic polynomial, we have [BH11, Section 1.4]

$$(-1)^{n-m}\Theta_\pi(g^*) = \Theta_{JL(\pi)}(g), \tag{3.5}$$

where Θ_π (resp. $\Theta_{JL(\pi)}$) is the character of π (resp. $JL(\pi)$).
For each representation $\pi \in \mathcal{A}^2_m(D)$, let

(i) $f(\pi)$ be the number of unramified characters χ of $F^\times$ that $\chi \otimes \pi \cong \pi$ (here χ is regarded as a representation of $G(F)$ by composing with the reduced norm map $\mathrm{Nrd} : G(F) \to F^\times$), and

(ii) $\delta(\pi)$ be the parametric degree of π (we do not require its full definition, so we only refer to [BH11, Section 2.7] for details).

It is known that $\delta(\pi)$ is a positive integer and is a multiple of $f(\pi)$. Moreover, π is supercuspidal if $\delta(\pi) = n$, while the converse is only true in the split case (when $G = G^*$).
Recall that we denote by $\mathcal{A}^0_n(F)$ the set of supercuspidal representations of $G^*(F)$. The correspondence (3.4) restricts to a bijection

$$JL : \mathcal{A}^0_n(F) \to \{\pi \in \mathcal{A}^2_m(D) | \delta(\pi) = n\}.$$

We call π *essentially tame* if p does not divide $\delta(\pi)/f(\pi)$. Let $\mathcal{A}^{\mathrm{et}}_m(D)$ be the set of isomorphism classes of irreducible representations in $\mathcal{A}^2_m(D)$ which are essentially tame and satisfy $\delta(\pi) = n$. Therefore $\mathcal{A}^{\mathrm{et}}_n(F)$ is the same collection defined in [BH05a]. Since the Jacquet-Langlands correspondence in (3.4) preserves the invariants $\delta(\pi)$ and $f(\pi)$, we have the following theorem [BH11, 2.8. Corollary 2].

THEOREM 3.1 (Essentially tame Jacquet-Langlands correspondence). *The restriction of the Jacquet-Langlands correspondence induces a bijection*

$$JL : \mathcal{A}^{\mathrm{et}}_n(F) \to \mathcal{A}^{\mathrm{et}}_m(D).$$

This bijection preserves the central characters on both sides. □

Bushnell and Henniart described this bijection explicitly in a way parallel to [BH05a], [BH05b], [BH10]. We recall the results briefly as follows. On the one hand, we have the bijection

$$_F\Pi : P_n(F) \to \mathcal{A}^{\mathrm{et}}_n(F),\ (E/F, \xi) \mapsto {}_F\Pi_\xi, \tag{3.6}$$

generalizing the construction of Howe [How77]. On the other hand, there is an analogous bijection

$$ {}_D\Pi : P_n(F) \to \mathcal{A}_m^{\mathrm{et}}(D),\ (E/F,\xi) \mapsto {}_D\Pi_\xi, \tag{3.7}$$

using the constructions in [Séc04], [Séc05a], [Séc05b], [SS08]. In fact, what is constructed in [BH11] is the inverse of (3.7), using the method called 'attached-pairs', since the construction parallel to (3.6) exhibits some 'novel technical difficulties' as mentioned in [BH11, Introduction 4.]. In the split case, the attached-pair method yields the inverse of (3.6) (see 4.4 of [BH11]).
The composition of the bijection in (3.6), the correspondence in Theorem 3.1, and the inverse of (3.7),

$$ {}_D\nu : P_n(F) \xrightarrow{{}_F\Pi} \mathcal{A}_n^{\mathrm{et}}(F) \xrightarrow{JL} \mathcal{A}_m^{\mathrm{et}}(D) \xrightarrow{{}_D\Pi^{-1}} P_n(F), \tag{3.8}$$

determines a tamely ramified quadratic character ${}_D\nu_\xi$ of $E^\times$ for each admissible character ξ of $E^\times$, depending only on the wild part of ξ, such that for each admissible pair $(E/F,\xi)$, the pair $(E/F, {}_D\nu_\xi\cdot\xi)$ is also admissible and

$$ {}_D\nu(E/F,\xi) = (E/F, {}_D\nu_\xi\cdot\xi). \tag{3.9}$$

We call this character ${}_D\nu_\xi$ the *rectifier* of ξ (for the Jacquet-Langlands correspondence). Using the First and Second Comparison Theorems of [BH11], we can compute the values of ${}_D\nu_\xi$. To express these values, we need the knowledge of certain invariants of finite symplectic modules, which will be described in Section 3.8. Finally, with the expression of ${}_D\nu_\xi$, we see that we can describe the correspondence in Theorem 3.1 explicitly, using (3.6), (3.7), and (3.8).

3.3 Some subgroups

We recall certain subgroups of $G(F)$. Suppose that the jump data $\{E_k, a_k | k = 0,\dots,t\}$ are defined by the factorization (3.1) of an admissible pair $(E/F,\xi)$, or equivalently, of its associated 1-pair $(E_0/F,\Xi)$. We require that $(E_0,\mathfrak{A})$ satisfy the conditions in Section 2.5. We write A_k the centralizer of E_k in A and $\mathfrak{A}_k = A_k\cap\mathfrak{A}$. We can then define $\mathfrak{P}_{\mathfrak{A}_k}$, $U_{\mathfrak{A}_k}$, $U^x_{\mathfrak{A}_k}$, and $U^{x+}_{\mathfrak{A}_k}$ for $x\in\mathbb{R}_{\geq 0}$ analogously as in Section 2.4. Following [Gra07, Definition 4.1], we construct the pro-p subgroups

$$\begin{aligned} H^1(\Xi,\mathfrak{A}) &= U^1_{\mathfrak{A}_0} U^{(a_0 e(\mathfrak{A}_1/\mathfrak{o}_E)/2)+}_{\mathfrak{A}_1}\cdots U^{(a_{t-1}e(\mathfrak{A}_t/\mathfrak{o}_E)/2)+}_{\mathfrak{A}_t} U^{(a_t e(\mathfrak{A}/\mathfrak{o}_E)/2)+}_{\mathfrak{A}} \text{ and}\\ J^1(\Xi,\mathfrak{A}) &= U^1_{\mathfrak{A}_0} U^{a_0 e(\mathfrak{A}_1/\mathfrak{o}_E)/2}_{\mathfrak{A}_1}\cdots U^{a_{t-1}e(\mathfrak{A}_t/\mathfrak{o}_E)/2}_{\mathfrak{A}_t} U^{a_t e(\mathfrak{A}/\mathfrak{o}_E)/2}_{\mathfrak{A}}. \end{aligned} \tag{3.10}$$

We also construct the subgroups

$$\begin{aligned} J(\Xi,\mathfrak{A}) &= U_{\mathfrak{A}_0} U^{a_0 e(\mathfrak{A}_1/\mathfrak{o}_E)/2}_{\mathfrak{A}_1}\cdots U^{a_{t-1}e(\mathfrak{A}_t/\mathfrak{o}_E)/2}_{\mathfrak{A}_t} U^{a_t e(\mathfrak{A}/\mathfrak{o}_E)/2}_{\mathfrak{A}} \text{ and}\\ \mathbf{J}(\Xi,\mathfrak{A}) &= E^\times J(\Xi,\mathfrak{A}) = E_0^\times J(\Xi,\mathfrak{A}). \end{aligned}$$

We abbreviate these groups by H^1, J^1, J and $\mathbf{J}$ if the admissible character ξ is fixed. Notice that H^1, J^1, J are compact subgroups of $G(F)$ and $\mathbf{J}$ is a compact-mod-center subgroup of $G(F)$.
In [Séc04, Section 3.1], these subgroups are defined based on a simple stratum $[\mathfrak{A}, -v_{\mathfrak{A}}(\beta), 0, \beta]$, where β is a suitable element in E_0, depending on Ξ and such that

$$-v_{\mathfrak{A}}(\beta) = \text{ the } E\text{-level of } \Xi \circ N_{E/E_0}.$$

The group $H^1(\Xi, \mathfrak{A})$ is denoted by $H^1(\beta, \mathfrak{A})$ in *loc. cit.* (and similarly for the other subgroups). This construction is an obvious generalization of [BK93, Section 3.1] (see also [BH11, Section 2.5] and the Comment therewithin).

3.4 Simple characters

Given an admissible 1-pair $(E_0/F, \Xi)$ and a finite unramified extension E/E_0, we define $H^1(\Xi, \mathfrak{A})$ as in (3.10). Using the idea of [Moy86, Section 3.2] (see also [Séc04, Definitions 3.22, 3.45, Proposition 3.47]), we attach to $(E_0/F, \Xi)$ a simple character $(H^1(\Xi, \mathfrak{A}), \theta_{\Xi,E})$ as follows. Suppose that Ξ admits a factorization of the form (3.3), with each $\xi_k \circ N_{E_0/E_k}$, where $k = 0, \dots, t+1$, a character of $U^1_{E_0}$. The generic conditions on the factorization imply that for each ξ_k there is $c_k \in E_k \cap \mathfrak{p}_E^{-a_k}$ such that $E_{k+1}[c_k] = E_k$ and

$$\xi_k \circ N_{E/E_k}(1+x) = \psi_F(\mathrm{tr}_{E/F}(c_k x)) \text{ for all } x \in \mathfrak{p}_E^{(a_k/2)+}. \tag{3.11}$$

Note that the element c_k can be chosen mod $\mathfrak{p}_E^{-a_k/2}$. We denote the character on the right side of (3.11) by ψ_{c_k}.
We define a character $\theta_{\Xi,E}$ of the subgroup $H^1(\Xi, \mathfrak{A})$ in (3.10) by the following inductive procedure. We first define a character θ_{t+1} on the subgroup $U_{\mathfrak{A}_{t+1}}^{(a_t e(\mathfrak{A}_{t+1}/\mathfrak{o}_E)/2)+}$ (note that indeed $E_{t+1} = F$ and $\mathfrak{A}_{t+1} = \mathfrak{A}$) by

$$\xi_{t+1} \circ \mathrm{Nrd}_{A/F} \text{ on } U_{\mathfrak{A}_{t+1}}^{(a_t e(\mathfrak{A}_{t+1}/\mathfrak{o}_E)/2)+}.$$

Inductively, suppose θ_{k+1} is defined, we construct θ_k on the subgroup

$$U_{\mathfrak{A}_k}^{(a_{k-1} e(\mathfrak{A}_k/\mathfrak{o}_E)/2)+} U_{\mathfrak{A}_{k+1}}^{(a_k e(\mathfrak{A}_{k+1}/\mathfrak{o}_E)/2)+} \cdots U_{\mathfrak{A}}^{(a_t e(\mathfrak{A}/\mathfrak{o}_E)/2)+}$$

by

$$\begin{aligned} &(\xi_k \circ \mathrm{Nrd}_{A_k/E_k}) \cdots (\xi_{t+1} \circ \mathrm{Nrd}_{A/F}) \text{ on } U_{\mathfrak{A}_k}^{(a_{k-1} e(\mathfrak{A}_k/\mathfrak{o}_E)/2)+} \text{ and} \\ &(\psi_{c_k} \circ \mathrm{trd}_{\mathfrak{A}_k/E_k})\theta_{k+1} \text{ on } U_{\mathfrak{A}_{k+1}}^{(a_k e(\mathfrak{A}_{k+1}/\mathfrak{o}_E)/2)+} \cdots U_{\mathfrak{A}}^{(a_t e(\mathfrak{A}/\mathfrak{o}_E)/2)+}. \end{aligned} \tag{3.12}$$

On the intersection $U_{\mathfrak{A}_k}^{(a_k e(\mathfrak{A}_k/\mathfrak{o}_E)/2)+}$, we have

$$\theta_{k+1} = (\xi_{k+1} \circ \mathrm{Nrd}_{A_{k+1}/E_{k+1}}) \cdots (\xi_{t+1} \circ \mathrm{Nrd}_{A/F})$$

and

$$\psi_{c_k} \circ \mathrm{trd}_{\mathfrak{A}_k/E_k} = \xi_k \circ \mathrm{Nrd}_{A_k/E_k}.$$

By Proposition 2.7 and (3.11), the two characters in (3.12) agree on the intersection, and so θ_k is defined. Finally, we define $\theta_{\Xi,E} = \theta_0$ on $H^1(\Xi, \mathfrak{A})$. By construction, this character $\theta_{\Xi,E}$ is normalized by $\mathfrak{K}_{\mathfrak{A}_0}$, hence it is a simple character by [Gra07, 5.3 Definition.(i)].
To get back our character Ξ from $\theta_{\Xi,E}$, just notice that $\theta_{\Xi,E}|_{U_E^1}$ factors through N_{E/E_0}, hence there is a unique character Ξ of $U_{E_0}^1$ such that $\Xi \circ N_{E/E_0} = \theta_{\Xi,E}|_{U_E^1}$.

3.5 Local constructions of attached pairs

We briefly summarize the construction in [BH11, Sec. 4] of the bijection ${}_D\Pi$ in (3.7). We will distinguish between the 'level-zero' case and the 'positive-level' case (we refer to [BH11, Sections 2.4 and 2.6] the definition of these cases). As we remarked in the Introduction, it is indeed the inverse of ${}_D\Pi$ that we are going to describe.
In the level-zero case, the construction is similar to the one in the split case. Each level zero $\pi \in \mathcal{A}_m^{\mathrm{et}}(D)$ contains a representation $(\mathrm{GL}_m(\mathfrak{o}_D), \lambda)$, called a maximal simple type of level zero, inflated from an irreducible cuspidal representation $(\mathrm{GL}_m(\mathbf{k}_D), \bar{\lambda})$. This representation $\bar{\lambda}$ corresponds, via Green's parametrization [Gre55], to a $\mathbf{k}_E/\mathbf{k}_F$-regular character $\bar{\xi}$ of $\mathbf{k}_E^\times$, where E/F is the unramified extension of degree n. We define the character ξ of $E^\times$ such that $\xi|_{\mathfrak{o}_E^\times}$ is the inflation of $\bar{\xi}$ and $\xi|_{F^\times}$ is the central character of π. By [BH11, 4.2. Proposition], the attached pair $(E/F, \xi)$ is admissible and the correspondence

$$\mathcal{A}_m^{\mathrm{et}}(D)_{level-0} \to P_n(F)_{level-0}, \, \pi \mapsto (E/F, \xi), \tag{3.13}$$

is bijective. We can show that $\pi = \mathrm{cInd}_{\mathbf{J}}^{G(F)} \Lambda$, where the condition $\delta(\pi) = n$ implies that $\mathbf{J} = F^\times \mathrm{GL}_m(\mathfrak{o}_D)$, and $(\mathbf{J}, \Lambda)$ is defined by the conditions

$$\Lambda|_{\mathrm{GL}_m(\mathfrak{o}_D)} = \lambda \text{ and } \Lambda|_{F^\times} \text{ is a multiple of } \xi|_{F^\times}. \tag{3.14}$$

The representation $(\mathbf{J}, \Lambda)$ is called an extended maximal simple type of level zero.
In the positive level case, we first recall the construction of a extended maximal simple type in general. Suppose we have a simple character (H^1, θ). For example, we can construct a simple character $\theta = \theta_{\Xi,E}$ as in Section 3.4 using an admissible pair $(E/F, \xi)$. We notice that the commutator subgroup $[J^1, J^1]$ lies in H^1 [BK93, (3.1.15)]. By [Séc04, Théorème 3.52], the above simple character θ induces an non-degenerate alternating bilinear form

$$h_\theta(x, y) = \theta([1 + x, 1 + y]), \text{ for all } 1 + x, 1 + y \in J^1, \tag{3.15}$$

on the $\mathbb{F}_p$-vector space

$${}_A\mathfrak{V}_\xi := J^1/H^1.$$

The classical theory of Heisenberg representation implies that there is a unique representation $\bar{\eta}$ of $J^1/\ker\theta$ containing the character θ of $H^1/\ker\theta$ as a central character. We then define η as the inflation of $\bar{\eta}$ to J^1. By [Séc05a, Théorème 2.28], there exists a unique irreducible representation (J, κ), which is called a β-extension (or wide extension in [BH11]) of η and satisfies certain conditions on its intertwining in $G(F)$ (see [BH11, (2.5.5)]). We now choose a maximal simple type $(\mathrm{GL}_{m_0}(\mathfrak{o}_{D_0}), \sigma)$ of $A_0^\times = Z_{G(F)}(E_0^\times)$ of level zero and inflate it to a representation (J, σ), since we know that $J = \mathrm{GL}_{m_0}(\mathfrak{o}_{D_0})J^1$. We obtain a maximal simple type (J, λ), where $\lambda = \kappa \otimes \sigma$. By [Séc05b, Théorème 5.2], there exists an irreducible representation Λ of $\mathbf{J} = {E_0}^\times J^1$ (by the condition $\delta(\pi) = n$), extending λ and whose compact-induction to $G(F)$ is irreducible and supercuspidal. The representation $(\mathbf{J}, \Lambda)$ is called an extended maximal simple type. By [BH11, Lemma 2 of Section 4.3], we can fix a unique extended type containing (J, λ) and satisfying the (non-canonical) conditions:

$$\varpi_F \in \ker \Lambda \text{ and } \det \Lambda \text{ has a } p\text{-power order.}$$

Following [BH11, Section 3 and 4], we have to approach indirectly to describe the inverse of ${}_D\Pi$. Suppose that we are given a representation $\pi \in \mathcal{A}_m^{\mathrm{et}}(D)$ of positive level. By [SS08, Théorème 5.21], it contains an extended maximal simple type $(\mathbf{J}, \Lambda)$ of the above form, such that $\Lambda|_{H^1}$ is a multiple of a simple character (H^1, θ). There is a unique character ξ_w of $E^\times$, depending on $\theta|_{U^1_{\mathfrak{A}_0}}$, satisfying the conditions in (3.2). In particular, we have

$$\theta|_{U^1_{\mathfrak{A}_0}} = \Xi \circ \mathrm{Nrd}_{A_0/E_0}, \text{ such that } \xi_w|_{U^1_E} = \Xi \circ N_{E/E_0}.$$

By the discussion of the previous paragraph, we can attach to ξ_w an extended maximal simple type $(\mathbf{J}, \Lambda_w)$ such that $\Lambda \cong \Lambda_{-1} \otimes \Lambda_w$ for a uniquely determined extended maximal simple type $(\mathbf{J}, \Lambda_{-1})$ of level zero. Attached to $(\mathbf{J}, \Lambda_{-1})$ is a level zero character ξ_{-1} of $E^\times$ admissible over E_0, as mentioned in the level zero case. Finally, by [BH11, 4.3. Proposition], the attached pair $(E/F, \xi)$, where $\xi = \xi_{-1}\xi_w$, is admissible and independent of the various choices above. We call $(E/F, \xi)$ a pair attached to π.
The technical part is to show that the attaching map

$$\mathcal{A}_m^{\mathrm{et}}(D) \to P_n(F),\ \pi \mapsto (E/F, \xi)$$

is well-defined and injective. This is done in the Parametrization Theorem of [BH11, Section 6]. The composition (3.8) is then injective (since the maps ${}_F\Pi$ and JL are known to be bijective) and preserves the restriction of each character to the subgroup $F^\times U_E^1$, which is of finite index of $E^\times$. Therefore, the map in (3.8) and hence ${}_D\Pi$ in (3.7), is bijective.

3.6 Finite symplectic modules

Since the group $\mathbf{J}$ normalizes the subgroups H^1, J^1, and the simple character θ of H^1, it acts on the finite quotient ${}_A\mathfrak{V}_\xi := J^1/H^1$. This quotient is denoted by ${}_{A^*}\mathfrak{V}_\xi$ in the split case $A = A^*$, which is studied in [BH10] and [Tam]. Notice that the quotient is clearly a finite dimensional $\mathbb{F}_p$-vector space. The action of $\mathbf{J}$ induces a symplectic $\mathbb{F}_p\mathbf{J}$-module structure on this space with respect to the non-degenerate alternating form h_θ in (3.15). We have a decomposition of

$$ {}_A\mathfrak{V}_\xi = {}_A\mathfrak{V}_{\xi,0} \oplus \cdots \oplus {}_A\mathfrak{V}_{\xi,t}, \tag{3.16}$$

into $\mathbb{F}_p\mathbf{J}$-submodules, where

$$ \begin{aligned} {}_A\mathfrak{V}_{\xi,k} &= \frac{U_{\mathfrak{A}_{k+1}}^{a_k e(\mathfrak{A}_{k+1}/\mathfrak{o}_E)/2}}{U_{\mathfrak{A}_k}^{a_k e(\mathfrak{A}_k/\mathfrak{o}_E)/2} U_{\mathfrak{A}_{k+1}}^{(a_k e(\mathfrak{A}_{k+1}/\mathfrak{o}_E)/2)+}} \\ &= \frac{\mathfrak{P}_{\mathfrak{A}_{k+1}}^{a_k e(\mathfrak{A}_{k+1}/\mathfrak{o}_E)/2}}{\mathfrak{P}_{\mathfrak{A}_k}^{a_k e(\mathfrak{A}_k/\mathfrak{o}_E)/2} + \mathfrak{P}_{\mathfrak{A}_{k+1}}^{(a_k e(\mathfrak{A}_{k+1}/\mathfrak{o}_E)/2)+}} \end{aligned} \tag{3.17}$$

for $k = 0, \dots, t$. In this paper, We call this decomposition the *coarse decomposition* of ${}_A\mathfrak{V}_\xi$. By [Séc04, Proposition 3.9], the decomposition (3.16) is orthogonal.

In the sequel, we will be interested in the adjoint action of $E^\times$ on ${}_A\mathfrak{V}_\xi$ restricted from that of $\mathbf{J}$, which factors through the finite group $\Psi_{E/F} := E^\times/F^\times(E^\times \cap J^1)$.

The following Proposition appears in [BH11, Proposition 5.6]. We re-interpret its proof here.

Proposition 3.2. *If E/F is totally ramified, then ${}_A\mathfrak{V}_\xi \cong {}_{A^*}\mathfrak{V}_\xi$.*

Proof. From the proofs of Propositions 4.1 and 4.2 (which are purely algebra and do not require the knowledge of this section), we see that the totally ramified condition implies that

$$ \mathfrak{P}_{\mathfrak{A}_k^*}^j / \mathfrak{P}_{\mathfrak{A}_k^*}^{j+1} \cong \mathfrak{P}_{\mathfrak{A}_k}^j / \mathfrak{P}_{\mathfrak{A}_k}^{j+1} \cong \mathrm{Ind}_1^{\Psi_{E/E_k}} \mathbf{k}_F \tag{3.18}$$

as a $\mathbf{k}_F\Psi_{E/E_k}$-module, for all $j \in \mathbb{Z}$ and $k = 1, \dots, t+1$. Moreover, we know that in the split case the index $e(\mathfrak{A}_k^*/\mathfrak{o}_E)$ (appearing in the powers in (3.17), when $A = A^*$) is always 1, and in general each $e(\mathfrak{A}_k/\mathfrak{o}_E)$ (again appearing in (3.17)) divides $f(E/E_k)$ (remember that $f(E/E_k)/e(\mathfrak{A}_k/\mathfrak{o}_E)$ is the integer s_{E_k} appearing in (2.7), when $K = E_k$), which is equal to 1 in the totally ramified case. Hence from the expression in (3.17) and using (3.18), we see that both ${}_A\mathfrak{V}_{\xi,k}$ and ${}_{A^*}\mathfrak{V}_{\xi,k}$ are isomorphic to

$$ \frac{\mathfrak{P}_{\mathfrak{A}_{k+1}}^{a_k/2}}{\mathfrak{P}_{\mathfrak{A}_k}^{a_k/2} + \mathfrak{P}_{\mathfrak{A}_{k+1}}^{(a_k/2)+}}. $$

Hence their sums ${}_A\mathfrak{V}_\xi$ and ${}_{A^*}\mathfrak{V}_\xi$ are also isomorphic. $\square$

3.7 Invariants of finite symplectic modules

Let $\mathbf{\Gamma}$ be a finite cyclic group whose order is not divisible by p. We call a finite $\mathbb{F}_p\mathbf{\Gamma}$-module V *symplectic* if there is a non-degenerate alternating form $\mathbf{h} : V \times V \to \mathbb{F}_p$ which is $\mathbf{\Gamma}$-invariant, in the sense that

$$\mathbf{h}({}^{\gamma}v_1, {}^{\gamma}v_2) = \mathbf{h}(v_1, v_2), \text{ for all } \gamma \in \mathbf{\Gamma},\ v_1, v_2 \in V.$$

The simple module V_λ corresponding to a character $\lambda \in \mathrm{Hom}(\mathbf{\Gamma}, \bar{\mathbb{F}}_p^\times)$ is the field $\mathbb{F}_p[\lambda(\mathbf{\Gamma})]$ generated over $\mathbb{F}_p$ by the image $\lambda(\mathbf{\Gamma})$, with $\mathbf{\Gamma}$-action

$${}^{\gamma}v = \lambda(\gamma)v, \text{ for all } \gamma \in \mathbf{\Gamma},\ v \in V_\lambda.$$

Its $\mathbb{F}_p$-linear dual $V_\lambda^* = \mathrm{Hom}(V_\lambda, \mathbb{F}_p)$ is isomorphic to $V_{\lambda^{-1}}$ by the map

$$V_{\lambda^{-1}} \to V_\lambda^*,\ v \mapsto (w \mapsto \mathrm{tr}_{\mathbb{F}_p[\lambda(\mathbf{\Gamma})]/\mathbb{F}_p}(wv)),$$

such that the canonical pairing $\langle\cdot,\cdot\rangle : V_\lambda \times V_{\lambda^{-1}} \to \mathbb{F}_p$ is $\mathbf{\Gamma}$-invariant.
We recall some basic facts from [BF83, (8.2.3)] and [BH10, Sec. 3, Prop. 4].

Proposition 3.3. *(i) An indecomposable symplectic $\mathbb{F}_p\mathbf{\Gamma}$-module is isomorphic to either one of the following two kinds,*

(a) a hyperbolic *module of the form $\mathbf{V}_\lambda = V_\lambda \oplus V_{\lambda^{-1}}$ such that either $\lambda^2 = 1$ or $V_\lambda \ncong V_{\lambda^{-1}}$, with the alternating form*

$$\mathbf{h}_{\mathbf{V}_\lambda}((v_1, v_1^*), (v_2, v_2^*)) = \langle v_1, v_2^*\rangle - \langle v_2, v_1^*\rangle,$$

for all $(v_1, v_1^*), (v_2, v_2^*) \in \mathbf{V}_\lambda$;

(b) an anisotropic *module of the form V_λ with $\lambda^2 \neq 1$ and $V_\lambda \cong V_{\lambda^{-1}}$. In this case, $[\mathbb{F}_p[\lambda(\mathbf{\Gamma})] : \mathbb{F}_p]$ is even and the alternating form $\mathbf{h}_{\mathbf{V}_\lambda}$ is defined, up to $\mathbf{\Gamma}$-isometry, by*

$$(v_1, v_2) \mapsto \mathrm{tr}_{\mathbb{F}_p[\lambda(\mathbf{\Gamma})]/\mathbb{F}_p}(\alpha v_1 \bar{v}_2), \text{ for all } v_1, v_2 \in V_\lambda,$$

where $v \mapsto \bar{v}$ is the $\mathbb{F}_p$-automorphism of $\mathbb{F}_p[\lambda(\mathbf{\Gamma})]$ of order 2 and $\alpha \in \mathbb{F}_p[\lambda(\mathbf{\Gamma})]^\times$ satisfies $\bar{\alpha} = -\alpha$.

(ii) If V_λ is anisotropic and $\mathbb{F}_p[\lambda(\Gamma)]_\pm$ denotes the subfield of $\mathbb{F}_p[\lambda(\Gamma)]$ such that $\mathbb{F}_p[\lambda(\Gamma)]/\mathbb{F}_p[\lambda(\Gamma)]_\pm$ is quadratic, then $\lambda(\mathbf{\Gamma})$ is a subgroup of $\ker(N_{\mathbb{F}_p[\lambda(\mathbf{\Gamma})]/\mathbb{F}_p[\lambda(\mathbf{\Gamma})]_\pm})$

(iii) The $\mathbf{\Gamma}$-isometry class of a symplectic $\mathbb{F}_p\mathbf{\Gamma}$-module $(V, \mathbf{h})$ is determined by the underlying $\mathbb{F}_p\mathbf{\Gamma}$-module V.

Part (iii) is particularly useful because, when we talk about invariants of $\mathbf{\Gamma}$-isometry classes of symplectic $\mathbb{F}_p\mathbf{\Gamma}$-modules, we do not have to write down the alternating forms explicitly.

Given a finite symplectic $\mathbb{F}_p\mathbf{\Gamma}$-module $\mathfrak{V}$, we attach a sign $t^0_{\mathbf{\Gamma}}(\mathfrak{V}) \in \{\pm 1\}$ and a quadratic character $t^1_{\mathbf{\Gamma}}(\mathfrak{V})$ of $\mathbf{\Gamma}$. We also set

$$t_{\mathbf{\Gamma}}(\mathfrak{V}) = t^0_{\mathbf{\Gamma}}(\mathfrak{V}) t^1_{\mathbf{\Gamma}}(\mathfrak{V})(\gamma),$$

where γ is any generator of $\mathbf{\Gamma}$. We call these *t-factors* of $\mathfrak{V}$.
We recall from [BH10, Section 3] the definition the t-factors.

(i) If $\mathbf{\Gamma}$ acts on $\mathfrak{V}$ trivially, then

$$t^0_{\mathbf{\Gamma}}(\mathfrak{V}) = 1 \text{ and } t^1_{\mathbf{\Gamma}}(\mathfrak{V}) \equiv 1.$$

(ii) Let $\mathfrak{V}$ be an indecomposable symplectic $\mathbb{F}_p\mathbf{\Gamma}$-module.

(a) If $\mathfrak{V} = \mathfrak{V}_\lambda \oplus \mathfrak{V}_{\lambda^{-1}}$ is hyperbolic, then

$$t^0_{\mathbf{\Gamma}}(\mathfrak{V}) = 1 \text{ and } t^1_{\mathbf{\Gamma}}(\mathfrak{V}) = \operatorname{sgn}_{\lambda(\mathbf{\Gamma})}(\mathfrak{V}_\lambda).$$

Here $\operatorname{sgn}_{\lambda(\mathbf{\Gamma})}(\mathfrak{V}_\lambda) : \mathbf{\Gamma} \to \{\pm 1\}$ is the character whose image $\gamma \mapsto \operatorname{sgn}_{\lambda(\gamma)}(\mathfrak{V}_\lambda)$ is the signature of the multiplication by $\lambda(\gamma)$ as a permutation of the set $\mathfrak{V}_\lambda$.

(b) If $\mathfrak{V} = \mathfrak{V}_\lambda$ is anisotropic, then

$$\begin{aligned} &t^0_{\mathbf{\Gamma}}(\mathfrak{V}) = -1 \text{ and} \\ &t^1_{\mathbf{\Gamma}}(\mathfrak{V})(\gamma) = \left(\frac{\gamma}{\ker(N_{\mathbb{F}_p[\lambda(\mathbf{\Gamma})]/\mathbb{F}_p[\lambda(\mathbf{\Gamma})]_\pm})}\right) \text{ for any } \gamma \in \mathbf{\Gamma}. \end{aligned}$$

Here $\left(\frac{\cdot}{\cdot}\right)$ is the symbol defined as follows: for every finite cyclic group H,

$$\left(\frac{x}{H}\right) = \begin{cases} 1 & \text{if } x \in H^2, \\ -1 & \text{otherwise.} \end{cases}$$

(iii) If $\mathfrak{V}$ decomposes into an orthogonal sum $\mathfrak{V}_1 \perp \cdots \perp \mathfrak{V}_t$ of indecomposable symplectic $\mathbb{F}_p\mathbf{\Gamma}$-modules, then

$$t^i_{\mathbf{\Gamma}}(\mathfrak{V}) = t^i_{\mathbf{\Gamma}}(\mathfrak{V}_1) \cdots t^i_{\mathbf{\Gamma}}(\mathfrak{V}_t) \text{ for } i = 0, 1.$$

Notice that when $p = 2$, the order of $\mathbf{\Gamma}$ is odd. In this case, $t^1_{\mathbf{\Gamma}}(\mathfrak{V})$ is always trivial, because all signature characters and symbols $\left(\frac{\cdot}{\cdot}\right)$ are trivial.

3.8 Values of rectifiers

Given a tamely ramified extension E/F and an F-admissible character ξ of $E^\times$, let ${}_D\nu_\xi$ be the rectifier of ξ defined in (3.9). To describe the values of ${}_D\nu_\xi$, we need to impose a condition on ϖ_E defined in (2.1):

$$\varpi_E \in E_0, \tag{3.19}$$

where E_0 be the first field appearing in the factorization (3.1) of ξ. This condition is the same as in the Second Comparison Theorem of [BH11, Section 7], where we further require that $\varpi_E^r \in F$ for some integer r coprime to p. Indeed, from the assumption in (2.1) this extra requirement is automatic in our situation. Under (3.19), the roots of unity $z_{E/F}$, z_{ϕ^i} (defined in Section 2.2), and others related to ϖ_E in later sections all depend on the first field E_0 in the jump data of ξ.
The values of the rectifier ${}_D\nu_\xi$ depends on the t-factors

$$t^1_{\boldsymbol{\mu}}({}_A\mathfrak{V}_\xi),\ t^1_{\boldsymbol{\mu}}({}_{A^*}\mathfrak{V}_\xi),\ t_{\langle\varpi\rangle}({}_A\mathfrak{V}_\xi) \text{ and } t_{\langle\varpi\rangle}({}_{A^*}\mathfrak{V}_\xi),$$

where $\boldsymbol{\mu} = \boldsymbol{\mu}_{E/F}$ and $\langle\varpi\rangle = \langle\varpi\rangle_{E/F}$ abbreviate the following subgroups of $\Psi_{E/F} = E^\times/F^\times U^1_E$,

$$\begin{aligned} &\boldsymbol{\mu} := \boldsymbol{\mu}_E/\boldsymbol{\mu}_F \text{ and}\\ &\langle\varpi\rangle := \text{the subgroup generated by the image of } \varpi_E. \end{aligned} \tag{3.20}$$

By the First and Second Comparison Theorems of [BH11], the rectifier ${}_D\nu_\xi$ has values

$$\begin{aligned} &{}_D\nu_\xi|_{\boldsymbol{\mu}_E} = t^1_{\boldsymbol{\mu}}({}_A\mathfrak{V}_\xi)t^1_{\boldsymbol{\mu}}({}_{A^*}\mathfrak{V}_\xi) \text{ and}\\ &{}_D\nu_\xi(\varpi_E) = (-1)^{n-m+f_\varpi-m_\varpi}t_{\langle\varpi\rangle}({}_A\mathfrak{V}_\xi)t_{\langle\varpi\rangle}({}_{A^*}\mathfrak{V}_\xi), \end{aligned} \tag{3.21}$$

where $f_\varpi := [E : F[\varpi_E]] = f(E/F[\varpi_E])$ and $m_\varpi = \gcd(m, f_\varpi)$.
In the case when E/F is totally ramified, Proposition 3.2 implies that

$${}_D\nu_\xi \text{ is unramified and } {}_D\nu_\xi(\varpi_E) = (-1)^{n-m},$$

as stated in [BH11, 5.3.Theorem].

4 Finite symplectic modules

4.1 Standard modules of central simple algebra

Let $\mathfrak{A}$ be the hereditary E-pure order in A, as discussed in Sections 2.4 and 2.5. The isomorphism (2.6) implies that $\mathfrak{P}^j_{\mathfrak{A}}/\mathfrak{P}^{j+1}_{\mathfrak{A}} \cong \mathrm{Mat}_s(\mathbf{k}_D)^r$ for all $j \in \mathbb{Z}$, where $s = \gcd(f, m)$ and $r = e/\gcd(d, e)$. We denote this quotient by $(\mathrm{Mat}_s(\mathbf{k}_D)^r)_j$ when we want to emphases the index j. Notice that as $\mathbf{k}_F\Psi_{E/F}$-modules, all $(\mathrm{Mat}_s(\mathbf{k}_D)^r)_j$, for j ranges over all $\mathbb{Z}$, are isomorphic to each other.
When $\mathfrak{A} = \mathfrak{A}^*$ and $j = 0$, we know that $\mathfrak{U}_{\mathfrak{A}^*} := \mathfrak{A}^*/\mathfrak{P}_{\mathfrak{A}^*} \cong \mathrm{Mat}_n(\mathbf{k}_F)$ admits a root-space decomposition

$$\mathfrak{U}_{\mathfrak{A}^*} \cong \mathfrak{U}_0 \bigoplus_{[\lambda]\in\Gamma_F\backslash\Phi} \mathfrak{U}_{[\lambda]},$$

where $\mathfrak{U}_0 \cong \mathfrak{o}_E/\mathfrak{p}_E$ on which $\Psi_{E/F}$ acts trivially, and $\mathfrak{U}_{[\lambda]}$ is the $\mathbf{k}_F$-subspace on which $\Psi_{E/F}$ acts by the character λ. Note that the equivalence class of the $\mathbf{k}_F\Psi_{E/F}$-module $\mathfrak{U}_{[\lambda]}$ depends only on the Γ_F-orbit of λ.

For future computation, we rewrite the above decomposition as

$$\mathfrak{U}_{\mathfrak{A}^*} \cong \bigoplus_{[g]\in\Gamma_E\backslash\Gamma_F/\Gamma_E} \mathfrak{U}_{[g]} \tag{4.1}$$

using the identification in Proposition 2.1. Here $\mathfrak{U}_{[g]} \cong \mathbf{k}_{E({}^gE)}$ as a $\mathbf{k}_F$-vector space for each $[g] \in \Gamma_E\backslash\Gamma_F/\Gamma_E$, and the $\Psi_{E/F}$-action on each $v \in \mathfrak{U}_{[g]}$ is given as follows: if $[g] = [\sigma^i\phi^j]$, then

$$ {}^zv = (z^{q^j-1})^{-1}v \text{ for all } z \in \boldsymbol{\mu}_E \text{ and } {}^{\varpi_E}v = (z_e^i z_{\phi^j})^{-1}v, \tag{4.2}$$

where z_e and z_{ϕ^j} are defined in Section 2.2.

For general inner form G, we first consider a simple case when A is a division algebra D. We write $(\mathbf{k}_D)_j := \mathfrak{P}_D^j/\mathfrak{P}_D^{j+1}$ for each $j \in \mathbb{Z}$.

Proposition 4.1. *For each $j \in \mathbb{Z}$, the $\mathbf{k}_F\Psi_{E/F}$-module $(\mathbf{k}_D)_j$ is isomorphic to*

$$\bigoplus_{[\sigma^i\phi^{hj}]\in\Gamma_E\backslash\Gamma_F/\Gamma_E} \mathfrak{U}_{[\sigma^i\phi^{hj}]}. \tag{4.3}$$

Proof. Recall that, if we denote by K the maximal unramified extension in D (of degree n over F), then

$$\mathfrak{P}_D^i = \cdots \oplus \mathfrak{p}_K\varpi_D^i \oplus \mathfrak{o}_K\varpi_D^{i+1} \oplus \cdots .$$

Hence we can use $x\varpi_D^i$, with $x \in \boldsymbol{\mu}_K \cup \{0\}$, as a representative in $\mathfrak{P}_D^i$ of an element in $(\mathbf{k}_D)_j$. We regard $(\mathbf{k}_D)_j$ as a $\mathbf{k}_E$-vector space of dimension e such that $z \in \boldsymbol{\mu}_E$ acts on each piece $(\mathbf{k}_E)_j$ by the character $\left[\begin{smallmatrix}1\\ \phi^{hj}\end{smallmatrix}\right](z) = z^{1-q^{hj}}$ as in (2.4). Therefore, a Frobenius reciprocity argument (which is still valid when p does not divide $\#\Psi_{E/F}$) implies that

$$(\mathbf{k}_D)_j \cong \mathrm{Ind}_{\boldsymbol{\mu}_E/\boldsymbol{\mu}_F}^{\Psi_{E/F}}(\mathbf{k}_E)_j;$$

more precisely, the action of ϖ_E has eigenvalues $(z_e^i z_{\phi^{hj}})^{-1}$, $i = 0,\dots,e-1$, where we recall that $z_{\phi^{hj}}$ is an eth root of $z_{E/F}^{q^{hj}-1}$. Hence we have the decomposition (4.3) with the $\Psi_{E/F}$-action on each component $\mathfrak{U}_{[\sigma^i\phi^{hj}]}$ as in (4.2) for each fixed j. $\square$

For the general A, if we write $(\mathrm{Mat}_s(\mathbf{k}_D)^r)_{j'} := \mathfrak{P}_{\mathfrak{A}}^{j'}/\mathfrak{P}_{\mathfrak{A}}^{j'+1}$ for each $j' \in \mathbb{Z}$, then we have the following result.

Proposition 4.2. *For each $j' \in \mathbb{Z}$, we have a decomposition*

$$(\mathrm{Mat}_s(\mathbf{k}_D)^r)_{j'} \cong \bigoplus_{\substack{[\sigma^i\phi^j]\in\Gamma_E\backslash\Gamma_F/\Gamma_E\\ j\equiv hj' \mod f/s}} \mathfrak{U}_{[\sigma^i\phi^j]}.$$

as a $\Psi_{E/F}$-module. (Recall that $f/s = d/\gcd(d,e) = e(\mathfrak{A}/\mathfrak{o}_E)$.)

Proof. Denote by $E_D = E \cap D$ the maximal subfield contained in both E and D, then $[\mathbf{k}_{E_D} : \mathbf{k}_F] = f/s$ and $[\mathbf{k}_D : \mathbf{k}_{E_D}] = \gcd(d,e) = e/r$. By Proposition 4.1, we know that $z \in \boldsymbol{\mu}_{E_D}$ acts on $(\mathrm{Mat}_s(\mathbf{k}_D)^r)_{j'}$ as a sum of $(\mathbf{k}_{E_D})_{hj'}$, i.e., z acts by the character $\left[\frac{1}{\bar{\phi}^{hj'}}\right](z) = z^{1-q^{hj'}}$, where $\bar{\phi}$ is the image of ϕ under the natural projection

$$\Gamma_{\mathbf{k}_E/\mathbf{k}_F} \to \Gamma_{\mathbf{k}_{E_D}/\mathbf{k}_F}. \tag{4.4}$$

(Note that the arguments above concerning Proposition 4.1 still valid even though E_D may not be a maximal subfield of D.) We now consider the $\mathbf{k}_{E_D}$-embeddings (where all choices are conjugate to each other)

$$\mathbf{k}_{E_D} \to \mathbf{k}_E \to \mathrm{Mat}_s(\mathbf{k}_{E_D}) \to \mathrm{Mat}_s(\mathbf{k}_D),$$

Notice that $\mathbf{k}_E$ is a maximal subfield of $\mathrm{Mat}_s(\mathbf{k}_{E_D})$. By the "twisted group ring decomposition", we know that $z \in \boldsymbol{\mu}_E$ acts on $\mathrm{Mat}_s(\mathbf{k}_{E_D})$ as a sum of $((\mathbf{k}_E)_{hj'})^s$, i.e., z acts on each of the s summands of $\mathbf{k}_E$ by the character $\left[\frac{1}{\phi^j}\right](z) = z^{1-q^j}$, for ϕ^j ranges over the s pre-images of $\bar{\phi}^{hj'}$ under the natural projection (4.4). We denote this $\boldsymbol{\mu}_E$-module by $\mathrm{Mat}_s(\mathbf{k}_{E_D})_{j'}$. Finally, since the relative degree of $(\mathrm{Mat}_s(\mathbf{k}_D))^r$ over $\mathrm{Mat}_s(\mathbf{k}_{E_D})$ is $e = [\Psi_{E/F} : \boldsymbol{\mu}_E/\boldsymbol{\mu}_F]$, a Frobenius reciprocity argument (which is still valid when p does not divide $\#\Psi_{E/F}$) implies that

$$(\mathrm{Mat}_s(\mathbf{k}_D)^r)_{j'} \cong \mathrm{Ind}_{\boldsymbol{\mu}_E/\boldsymbol{\mu}_F}^{\Psi_{E/F}} \mathrm{Mat}_s(\mathbf{k}_{E_D})_{j'};$$

Therefore, we have obtained the desired decomposition and proved the proposition. □

The following Corollary is a direct consequence of Proposition 4.2.

Corollary 4.3. *The graded algebra*

$$\mathfrak{U}_{\mathfrak{A}} := \bigoplus_{j'=0}^{f/s-1} (\mathrm{Mat}_s(\mathbf{k}_D)^r)_{j'},$$

is isomorphic to $\mathfrak{U}_{\mathfrak{A}^*}$ as a $\Psi_{E/F}$-module. □

We provide some notations for later use. We write

$$\mathfrak{U}_{\mathrm{sym}} := \bigoplus_{[g] \in (\Gamma_E \backslash \Gamma_F / \Gamma_E)_{\mathrm{sym}}} \mathfrak{U}_{[g]}$$

and also $\mathfrak{U}_{\mathrm{sym-ram}}$ and $\mathfrak{U}_{\mathrm{sym-unram}}$ analogously. Given intermediate extensions $F \subseteq K \subseteq L \subseteq E$, we write

$$\mathfrak{U}_{K/L} := \bigoplus_{[g] \in \Gamma_E \backslash (\Gamma_L - \Gamma_K) / \Gamma_E} \mathfrak{U}_{[g]}$$

We also define the *symmetric module* associated to $\mathfrak{U}_{[g]}$ (or $\mathfrak{U}_{[g^{-1}]}$) by

$$\boldsymbol{\mathfrak{U}}_{[g]} := \begin{cases} \mathfrak{U}_{[g]} \oplus \mathfrak{U}_{[g^{-1}]} & \text{if } [g] \text{ is asymmetric,} \\ \mathfrak{U}_{[g]} & \text{if } [g] \text{ is symmetric,} \end{cases} \tag{4.5}$$

and call

$$\mathfrak{U}_{\mathfrak{A}} \cong \mathfrak{U}_0 \bigoplus_{[g]\in(\Gamma_E\backslash\Gamma_F/\Gamma_E)_{\mathrm{sym}}\sqcup(\Gamma_E\backslash\Gamma_F/\Gamma_E)_{\mathrm{asym}/\pm}} \boldsymbol{\mathfrak{U}}_{[g]} \tag{4.6}$$

the *complete symmetric decomposition* of $\mathfrak{U}_{\mathfrak{A}}$. If $\mathfrak{V}$ is a submodule of $\mathfrak{U}_{\mathfrak{A}}$, we also use the same convention to denote its submodules, for example, $\mathfrak{V}_{K/L} = \mathfrak{U}_{K/L} \cap \mathfrak{V}$ and $\boldsymbol{\mathfrak{V}}_{[g]} = \boldsymbol{\mathfrak{U}}_{[g]} \cap \mathfrak{V}$, and also call

$$\mathfrak{V} \cong (\mathfrak{U}_0 \cap \mathfrak{V}) \bigoplus_{[g]\in(\Gamma_E\backslash\Gamma_F/\Gamma_E)_{\mathrm{sym}}\sqcup(\Gamma_E\backslash\Gamma_F/\Gamma_E)_{\mathrm{asym}/\pm}} \boldsymbol{\mathfrak{V}}_{[g]}$$

the complete symmetric decomposition of $\mathfrak{V}$

4.2 Complete decomposition of finite symplectic modules

We are interested in the adjoint action of $E^\times$ on ${}_A\mathfrak{V}_\xi$ restricted from that of $\mathbf{J}$, which factors through the finite group $E^\times/F^\times(E^\times \cap J^1) \cong \Psi_{E/F}$. We also know that this action preserves the symplectic structure h_θ (3.15) on ${}_A\mathfrak{V}_\xi$. Hence ${}_A\mathfrak{V}_\xi$ is moreover a finite symplectic $\mathbb{F}_p\Gamma$-module for each cyclic subgroup Γ of $\Psi_{E/F}$. We denote the $\mathfrak{U}_{[g]}$-isotypic component in ${}_A\mathfrak{V}_\xi$ by ${}_A\mathfrak{V}_{\xi,[g]}$, and obtain the decompositions

$${}_A\mathfrak{V}_\xi = \bigoplus_{[g]\in(\Gamma_E\backslash\Gamma_F/\Gamma_E)'} {}_A\mathfrak{V}_{\xi,[g]} = \bigoplus_{[g]\in(\Gamma_E\backslash\Gamma_F/\Gamma_E)_{\mathrm{sym}}\sqcup(\Gamma_E\backslash\Gamma_F/\Gamma_E)_{\mathrm{asym}/\pm}} {}_A\boldsymbol{\mathfrak{V}}_{\xi,[g]} \tag{4.7}$$

inherited from (4.1) and (4.6) respectively. These decomposition are finer than the one in (3.16). Indeed, it is easy to see that

$${}_A\mathfrak{V}_{\xi,k} := {}_A\mathfrak{V}_{\xi,E_k/E_{k+1}} = \bigoplus_{[g]\in\Gamma_E\backslash(\Gamma_{E_{k+1}}-\Gamma_{E_k})/\Gamma_E} {}_A\mathfrak{V}_{\xi,[g]}$$

for $k = 0, \ldots, t$.

Proposition 4.4. *The complete symmetric decomposition of ${}_A\mathfrak{V}_\xi$ is orthogonal with respect to the alternating form h_θ.*

Proof. Since we know that the $\Psi_{E/F}$-components of ${}_A\mathfrak{V}_\xi$ consist of those in the standard module $\mathfrak{U}_{\mathfrak{A}}$, which is isomorphic to the standard one $\mathfrak{U}_{\mathfrak{A}^*}$ in the split case, the proof of the assertion is just analogous to the one in the split case [Tam, (5.10)], based on the argument of [BF83, (8.2.3),(8.2.4)]. □

We would like to describe the isotypic component appearing in the complete decomposition (4.7) of ${}_A\mathfrak{V}_\xi$. We first write $e(\mathfrak{A}/\mathfrak{A}_{k+1}) := e(\mathfrak{A}/\mathfrak{o}_E)/e(\mathfrak{A}_{k+1}/\mathfrak{o}_E)$.

Proposition 4.5. *The quotient $\mathfrak{P}^{j}_{\mathfrak{A}} \cap A_{k+1}/\mathfrak{P}^{j+}_{\mathfrak{A}} \cap A_{k+1}$ is non-trivial if and only if $j \in e(\mathfrak{A}/\mathfrak{A}_{k+1})\mathbb{Z}$.*

Proof. Since $\mathfrak{P}^{j}_{\mathfrak{A}} \cap A_{k+1} = \mathfrak{P}^{j/e(\mathfrak{A}/\mathfrak{A}_{k+1})}_{\mathfrak{A}_{k+1}}$ for all $j \in \mathbb{Z}$, the assertion follows directly. $\square$

We now specify $j = j_k = e(\mathfrak{A}/\mathfrak{o}_E)a_k/2$ for some integer a_k, and so $\mathfrak{P}^{j_k}_{\mathfrak{A}} \cap A_{k+1} = \mathfrak{P}^{a_k e(\mathfrak{A}_{k+1}/\mathfrak{o}_E)/2}_{\mathfrak{A}_{k+1}}$, such that the index on the right side is the one appearing in the group J^1 (3.10). The condition in Proposition 4.5 is satisfied if and only if a_k is even or $e(\mathfrak{A}_{k+1}/\mathfrak{o}_E)$ is even, in which case

$$
\begin{aligned}
{}_A\mathfrak{V}_{\xi,k} &\cong \frac{\mathfrak{P}^{a_k e(\mathfrak{A}_{k+1}/\mathfrak{o}_E)/2}_{\mathfrak{A}_{k+1}}}{\mathfrak{P}^{a_k e(\mathfrak{A}_{k+1}/\mathfrak{o}_E)/2+}_{\mathfrak{A}_{k+1}} + \mathfrak{P}^{a_k e(\mathfrak{A}_k/\mathfrak{o}_E)/2}_{\mathfrak{A}_k}} \\
&\cong \begin{cases} \mathrm{Mat}_{s_{k+1}}(\mathbf{k}_{D_{k+1}})^{r_{k+1}}/\mathrm{Mat}_{s_k}(\mathbf{k}_{D_k})^{r_k} & \text{when } a_k e(\mathfrak{A}_k/\mathfrak{o}_E)/2 \in \mathbb{Z}, \\ \mathrm{Mat}_{s_{k+1}}(\mathbf{k}_{D_{k+1}})^{r_{k+1}} & \text{otherwise,} \end{cases}
\end{aligned}
$$

where r_k and s_k are the invariants of $\mathfrak{A}_k$ analogous to r and s of $\mathfrak{A}$.
To summarize, ${}_A\mathfrak{V}_{\xi,k}$ is isomorphic to

0 when a_k is odd and $e(\mathfrak{A}_{k+1}/\mathfrak{o}_E)$ is odd,

$\mathrm{Mat}_{s_{k+1}}(\mathbf{k}_{D_{k+1}})^{r_{k+1}}$
when a_k is odd, $e(\mathfrak{A}_k/\mathfrak{o}_E)$ is odd, and $e(\mathfrak{A}_{k+1}/\mathfrak{o}_E)$ is even,

$\mathrm{Mat}_{s_{k+1}}(\mathbf{k}_{D_{k+1}})^{r_{k+1}}/\mathrm{Mat}_{s_k}(\mathbf{k}_{D_k})^{r_k}$
when a_k is even or $e(\mathfrak{A}_k/\mathfrak{o}_E)$ is even.

The action of $\Psi_{E/F}$ on $\mathfrak{P}^{j_k}_{\mathfrak{A}} \cap A_{k+1}$ is given by $\sigma^i\phi^{jh} \in \Gamma_{E_{k+1}}$, where j has image hj_k in the natural projection $\mathbb{Z}/f\mathbb{Z} \to \mathbb{Z}/e(\mathfrak{A}/\mathfrak{o}_E)\mathbb{Z}$. Therefore, directly from the description of ${}_A\mathfrak{V}_{\xi,k}$ above, we have the following decompositions.

Proposition 4.6. *The complete decomposition of the component ${}_A\mathfrak{V}_{\xi,k}$ is given as follows.*

(i) When a_k is odd and $e(\mathfrak{A}_{k+1}/\mathfrak{o}_E)$ is odd, then ${}_A\mathfrak{V}_{\xi,k}$ is trivial.

(ii) When a_k is odd, $e(\mathfrak{A}_k/\mathfrak{o}_E)$ is odd, and $e(\mathfrak{A}_{k+1}/\mathfrak{o}_E)$ is even, then

$$
{}_A\mathfrak{V}_{\xi,k} \cong \bigoplus_{\substack{[g]=[\sigma^i\phi^j]\in\Gamma_E\backslash\Gamma_{E_{k+1}}/\Gamma_E \\ j\equiv hj_k \mod e(\mathfrak{A}/\mathfrak{o}_E)}} \mathfrak{U}_{[g]}.
$$

(iii) When a_k is even or $e(\mathfrak{A}_k/\mathfrak{o}_E)$ is even, then

$$
{}_A\mathfrak{V}_{\xi,k} \cong \bigoplus_{\substack{[g]=[\sigma^i\phi^j]\in\Gamma_E\backslash(\Gamma_{E_{k+1}}-\Gamma_{E_k})/\Gamma_E \\ j\equiv hj_k \mod e(\mathfrak{A}/\mathfrak{o}_E)}} \mathfrak{U}_{[g]}.
$$

$\square$

4.3 Some properties of parities of jumps

Let R be the index when $f(E/E_R)$ is odd and $f(E/E_{R+1})$ is even.

Lemma 4.7. *We have* $(\Gamma_{E_{R+1}}/\Gamma_E)_{\mathrm{sym-unram}} = ((\Gamma_{E_{R+1}} - \Gamma_{E_R})/\Gamma_E)_{\mathrm{sym-unram}}$.

Proof. Recall Proposition 2.2 that every symmetric unramified $[g]$ are of the form $[\sigma^i \phi^{f/2}]$, so there is no coset of the form $\sigma^i \phi^{f/2}$ belonging to Γ_{E_R}. $\square$

Let Q be the index when $e(\mathfrak{A}_Q/\mathfrak{o}_E)$ is odd and $e(\mathfrak{A}_{Q+1}/\mathfrak{o}_E)$ is even.

Lemma 4.8. *Suppose that f is even. We always have $R \leq Q$. If moreover m is odd, then $Q = R$.*

Proof. We know that $e(\mathfrak{A}_k/\mathfrak{o}_E)$ divides $f(E/E_k)$, so that if $Q < R$, then the even number $e(\mathfrak{A}_{Q+1}/\mathfrak{o}_E)$ divides $f(E/E_{Q+1})$, which divides the odd number $f(E/E_R)$. This is a contradiction. Hence $R \leq Q$. When $R \lneq Q$, then

$$e(\mathfrak{A}_{R+1}/\mathfrak{o}_E) \text{ is odd and } f(E/E_{R+1}) \text{ is even.} \tag{4.8}$$

Since

$$e(\mathfrak{A}_{R+1}/\mathfrak{o}_E) = f_{E_{R+1}}/s_{E_{R+1}} = \frac{f(E/E_{R+1})}{\gcd(f(E/E_{R+1}), \gcd(m, n(E/E_{R+1})))}, \tag{4.9}$$

the statement (4.8) is equivalent to saying that

$$\begin{array}{l}\text{the 2-powers of the numerator and}\\ \text{the denominator on the right side of (4.9) are equal.}\end{array} \tag{4.10}$$

This power is greater than 0. Hence (4.10) is equivalent to that

$$(\text{the 2-power of } m) \geq (\text{the 2-power of } f(E/E_{R+1})) \gneq 0. \tag{4.11}$$

If m is odd, then (4.11) is a contradiction. $\square$

4.4 Symmetric submodules

We write ${}_A\mathfrak{V}_{\xi,\mathrm{sym}} = {}_A\mathfrak{V}_\xi \cap \mathfrak{U}_{\mathrm{sym}}$ and ${}_A\mathfrak{V}_{\xi,\mathrm{sym-ram}}$ and ${}_A\mathfrak{V}_{\xi,\mathrm{sym-unram}}$ analogously.

4.4.1 Case when f is odd

From Proposition 3.2, we always have

$${}_A\mathfrak{V}_{\xi,\mathrm{sym}} = {}_A\mathfrak{V}_{\xi,\mathrm{sym-ram}} \cong {}_{A^*}\mathfrak{V}_{\xi,\mathrm{sym-ram}} = {}_{A^*}\mathfrak{V}_{\xi,\mathrm{sym}}. \tag{4.12}$$

4.4.2 Case when f is even

Notice that the natural projection $\mathbb{Z}/f\mathbb{Z} \to \mathbb{Z}/e(\mathfrak{A}/\mathfrak{o}_E)\mathbb{Z}$ maps

$$f/2 \mapsto \begin{cases} 0 & \text{if } e(\mathfrak{A}/\mathfrak{o}_E) \text{ divides } f/2, \\ e(\mathfrak{A}/\mathfrak{o}_E)/2 \neq 0 & \text{otherwise} \end{cases} \tag{4.13}$$

The condition that $e(\mathfrak{A}/\mathfrak{o}_E)$ divides $f/2$ is equivalent to that s is even. When f is even, then $s = \gcd(f, m)$ is even if and only if m is even. We hence separate the cases according to the parity of m.

4.4.3 Case when both f and m are even

In this case, $f/2$ is mapped to 0 by $\mathbb{Z}/f\mathbb{Z} \to \mathbb{Z}/e(\mathfrak{A}/\mathfrak{o}_E)\mathbb{Z}$. We separate the cases according to the parity of the jump a_k. When a_k is odd, neither 0 or $f/2$ is mapped to $hj_k \neq 0 \in \mathbb{Z}/e(\mathfrak{A}/\mathfrak{o}_E)\mathbb{Z}$, and so ${}_A\mathfrak{V}_{\xi,k}$ is trivial. When a_k is even, both 0 and $f/2$ are mapped to $hj_k = 0$ by (4.13), and so ${}_A\mathfrak{V}_{\xi,k,\mathrm{sym}} \cong \mathfrak{U}_{k,\mathrm{sym}}$. We also recall that

$${}_{A^*}\mathfrak{V}_{\xi,k} \cong \begin{cases} 0 & \text{if } a_k \text{ is odd}, \\ \mathfrak{U}_{k,\mathrm{sym}} & \text{if } a_k \text{ is even}. \end{cases}$$

Whatever the parity of a_k is, we always have ${}_A\mathfrak{V}_{\xi,\mathrm{sym}} \cong {}_{A^*}\mathfrak{V}_{\xi,\mathrm{sym}}$.

4.4.4 Case when f is even and m is odd

In this case, notice that $e(\mathfrak{A}/\mathfrak{o}_E)$ must be even, and

$$j_k = e(\mathfrak{A}/\mathfrak{o}_E)a_k/2 \equiv \begin{cases} 0 \\ e(\mathfrak{A}/\mathfrak{o}_E)/2 \end{cases} \mod e(\mathfrak{A}/\mathfrak{o}_E) \begin{cases} \text{if } a_k \text{ is even}, \\ \text{if } a_k \text{ is odd}. \end{cases}$$

Therefore,

$${}_A\mathfrak{V}_{\xi,k,\mathrm{sym}} = \begin{cases} {}_A\mathfrak{V}_{\xi,k,\mathrm{sym-ram}} & \text{if } a_k \text{ is even}, \\ {}_A\mathfrak{V}_{\xi,k,\mathrm{sym-unram}} & \text{if } a_k \text{ is odd}. \end{cases}$$

Using Proposition 4.6, we find that when a_k is even,

$${}_A\mathfrak{V}_{\xi,k,\mathrm{sym}} = \mathfrak{U}_{E_k/E_{k+1},\mathrm{sym-ram}} = \bigoplus_{[\sigma^i] \in (\Gamma_E \backslash (\Gamma_{E_{R+1}} - \Gamma_{E_R})/\Gamma_E)_{\mathrm{sym}}} \mathfrak{U}_{[\sigma^i]},$$

Here the index R is defined in Section 4.3. When a_k is odd, ${}_A\mathfrak{V}_{\xi,k,\mathrm{sym}}$ is trivial when $k < R$, and is isomorphic to

$$\mathfrak{U}_{E/E_{R+1},\mathrm{sym-unram}} = \bigoplus_{[\sigma^i\phi^{f/2}] \in (\Gamma_E \backslash \Gamma_{E_R}/\Gamma_E)_{\mathrm{sym}}} \mathfrak{U}_{[\sigma^i\phi^{f/2}]}$$

when $k = R$, and to

$$\mathfrak{U}_{E_k/E_{k+1},\mathrm{sym-unram}} = \bigoplus_{[\sigma^i\phi^{f/2}]\in(\Gamma_E\backslash(\Gamma_{E_{k+1}}-\Gamma_{E_k})/\Gamma_E)_{\mathrm{sym}}} \mathfrak{U}_{[\sigma^i\phi^{f/2}]}$$

when $k > R$.

We observe that, whether a_k is odd or even, the symmetric unramified part of ${}_A\mathfrak{V}_\xi$ and ${}_{A^*}\mathfrak{V}_\xi$ are complementary, in the sense that

$${}_A\mathfrak{V}_{\xi,k,\mathrm{sym-unram}} \oplus {}_{A^*}\mathfrak{V}_{\xi,k,\mathrm{sym-unram}} = \mathfrak{U}_{E_k/E_{k+1},\mathrm{sym-unram}}$$

for all $k = 0, \dots, t$

We summarize the above in the following.

Proposition 4.9. *We always have* ${}_A\mathfrak{V}_{\xi,\mathrm{sym-ram}} \cong {}_{A^*}\mathfrak{V}_{\xi,\mathrm{sym-ram}}$ *and*

(i) *when f is odd, or when both f and m are even, then* ${}_A\mathfrak{V}_{\xi,\mathrm{sym-unram}} \cong {}_{A^*}\mathfrak{V}_{\xi,\mathrm{sym-unram}}$*;*

(ii) *when f is even and m is odd, then* ${}_A\mathfrak{V}_{\xi,\mathrm{sym-unram}} \oplus {}_{A^*}\mathfrak{V}_{\xi,\mathrm{sym-unram}} = \mathfrak{U}_{\mathrm{sym-unram}}$.

4.5 t-factors of isotypic components

We recall the values of the t-factors $t^i_{\mathbf{\Gamma}}(\mathfrak{V})$, $i = 0, 1$, when $\mathbf{\Gamma}$ is one of the cyclic subgroups $\boldsymbol{\mu}$ and $\langle\varpi\rangle$ of $\Psi_{E/F}$ defined in (3.20), and $\mathfrak{V}$ is a symmetric module $\mathfrak{U}_{[g]}$ defined in (4.5). The following Proposition describes all $t^i_{\mathbf{\Gamma}}(\mathfrak{U}_{[g]})$ except when $[g] = [\sigma^{e/2}]$.

Proposition 4.10 ([Tam, Proposition 4.9]). (i) *If $[g] = [\sigma^i\phi^j]$ is asymmetric, then*

$$t^0_{\boldsymbol{\mu}}(\mathfrak{U}_{[g]}) = 1, \quad t^1_{\boldsymbol{\mu}}(\mathfrak{U}_{[g]}) : z \mapsto \mathrm{sgn}_{z^{q^i-1}}(\mathfrak{U}_{[g]}),$$
$$t^0_{\langle\varpi\rangle}(\mathfrak{U}_{[g]}) = 1, \ \textit{and } t^1_{\langle\varpi\rangle}(\mathfrak{U}_{[g]})(\varpi_E) = \mathrm{sgn}_{z_e^i z_{\phi^j}}(\mathfrak{U}_{[g]}).$$

(ii) *If $[g] = [\sigma^i]$ is symmetric and not equal to $[1]$ or $[\sigma^{e/2}]$, then*

$$t^0_{\boldsymbol{\mu}}(\mathfrak{U}_{[g]}) = 1, \quad t^1_{\boldsymbol{\mu}}(\mathfrak{U}_{[g]}) \equiv 1,$$
$$t^0_{\langle\varpi\rangle}(\mathfrak{U}_{[g]}) = -1, \ \textit{and } t^1_{\langle\varpi\rangle}(\mathfrak{U}_{[g]}) : \varpi_E \mapsto \left(\frac{z_e^i}{\ker(N_{\mathbb{F}_p[z_e^i]/\mathbb{F}_p[z_e^i]_\pm})}\right).$$

(iii) *If $[g] = [\sigma^i\phi^{f/2}]$ is symmetric, then*

$$t^0_{\boldsymbol{\mu}}(\mathfrak{U}_{[g]}) = -1, \quad t^1_{\boldsymbol{\mu}}(\mathfrak{U}_{[g]}) \ \textit{is quadratic},$$

and

(I) *if* $z_e^i z_{\phi^{f/2}} = 1$, *then* $t^0_{\langle\varpi\rangle}(\mathfrak{U}_{[g]}) = 1$ *and* $t^1_{\langle\varpi\rangle}(\mathfrak{U}_{[g]}) \equiv 1$*;*

(II) *if* $z_e^i z_{\phi^{f/2}} = -1$, *then* $t^0_{\langle\varpi\rangle}(\mathfrak{U}_{[g]}) = 1$ *and* $t^1_{\langle\varpi\rangle}(\mathfrak{U}_{[g]})(\varpi_E) = (-1)^{\frac{1}{2}(q^{f/2}-1)}$*;*

(III) *if* $z_e^i z_{\phi^{f/2}} \neq \pm 1$, *then* $t^0_{\langle\varpi\rangle}(\mathfrak{U}_{[g]}) = -1$ *and*

$$t^1_{\langle\varpi\rangle}(\mathfrak{U}_{[g]}) : \varpi_E \mapsto \left(\frac{z_e^i z_{\phi^{f/2}}}{\ker(N_{\mathbb{F}_p[z_e^i z_{\phi^{f/2}}]/\mathbb{F}_p[z_e^i z_{\phi^{f/2}}]_\pm})}\right).$$

□

In the exceptional case, when $[g] = [\sigma^{e/2}]$, we have $\boldsymbol{\mu}_{E_g} = \boldsymbol{\mu}_E$. To unify notation, we define

$$t^1_{\boldsymbol{\mu}}(\mathfrak{U}_{[\sigma^{e/2}]}) : \boldsymbol{\mu}_E \to \{\pm 1\},\, x \mapsto \left(\frac{x}{\boldsymbol{\mu}_E}\right). \tag{4.14}$$

The $\mathbb{F}_p\langle\varpi\rangle$-module structure of $\mathfrak{U}_{[\sigma^{e/2}]}$ does not concern us (see the explanation after Formula (5.5)).

The following properties concerning symmetric double cosets are useful when computing the above t-factors.

Proposition 4.11. *Suppose that* $[g]$ *is symmetric.*

(i) *If* $[g]$ *is ramified (resp. unramified), then* $[\mathfrak{U}_{[g]} : \mathbf{k}_E]$ *is even (resp. odd).*

(ii) *Let* $\mathbb{F}_p[\left[\frac{1}{g}\right](\varpi_E)]$ *be the field extension of* $\mathbb{F}_p$ *generated by the image of* $\left[\frac{1}{g}\right](\varpi_E)$ *in* $\bar{\mathbf{k}}_F^\times$. *If* $[g] \neq [\sigma^{e/2}]$, *then the degree* $[\mathfrak{U}_{[g]} : \mathbb{F}_p[\left[\frac{1}{g}\right](\varpi_E)]]$ *is odd.*

Proof. The first statement for ramified $[g]$ is a simple calculation, and that for unramified $[g]$ is a consequence of Proposition 2.3. The second statement is proved in [Tam, Lemma 4.8]. □

We would like to extend our definition of the t-factors $t^i_{\boldsymbol{\mu}}(\mathfrak{U}_{[g]})$, with $i = 0, 1$, from $\boldsymbol{\mu}$ to $\boldsymbol{\mu}_g = \boldsymbol{\mu}_{E_g}/\boldsymbol{\mu}_F$. We define

$$t^0_{\boldsymbol{\mu}_g}(\mathfrak{U}_{[g]}) := t^0_{\boldsymbol{\mu}}(\mathfrak{U}_{[g]})$$

and for all $z \in \boldsymbol{\mu}_g$,

$$t^1_{\boldsymbol{\mu}_g}(\mathfrak{U}_{[g]}) : z \mapsto \begin{cases} \mathrm{sgn}_{\left[\frac{1}{g}\right](z)}(\mathfrak{U}_{[g]}) & \text{if } [g] \text{ is asymmetric,} \\ \left(\dfrac{\left[\frac{1}{\sigma^i \phi^{f[\mathfrak{U}_{[g]}:\mathbf{k}_E]/2}}\right](z)}{\ker N_{\mathbf{k}_{E_g}/\mathbf{k}_{E_{\pm g}}}}\right) & \text{if } [g] \text{ is symmetric.} \end{cases}$$

Proposition 4.12. *The restriction* $t^1_{\boldsymbol{\mu}_g}(\mathfrak{U}_{[g]})$ *to* $\boldsymbol{\mu}$ *is* $t^1_{\boldsymbol{\mu}}(\mathfrak{U}_{[g]})$.

Proof. For asymmetric $[g]$, the result is immediate by definition. For symmetric ramified $[g]$, the restriction of the root $\left[\frac{1}{\sigma^i \phi^{f[\mathfrak{U}_{[g]}:\mathbf{k}_E]}}\right]$ to $\boldsymbol{\mu} = \boldsymbol{\mu}_E/\boldsymbol{\mu}_F$ is trivial, so the assertion is again true. When $[g]$ is symmetric unramified, we have to show that the restriction of $\left(\frac{\cdot}{\ker N_{\mathbf{k}_{E_g}/\mathbf{k}_{E_{\pm g}}}}\right)$ to $\boldsymbol{\mu}$ is $\left(\frac{\cdot}{\ker N_{\mathbf{k}_E/\mathbf{k}_{E\pm}}}\right)$, or equivalently, to show that the index of the subgroup $\ker N_{\mathbf{k}_E/\mathbf{k}_{E\pm}} \cong \boldsymbol{\mu}_{q^{f/2}+1}$ of $\ker N_{\mathbf{k}_E/\mathbf{k}_{E\pm}} \cong \boldsymbol{\mu}_{q^{f[\mathfrak{U}_{[g]}:\mathbf{k}_E]/2}+1}$ is odd, which follows from Proposition 4.11. □

5 Zeta-data

5.1 Admissible embeddings of L-tori

As mentioned in Section 1.3, to understand ζ-data, it is better to first understand χ-data, which is motivated by constructing admissible embeddings of L-tori [LS87, Section 2.6].

We take T to be an elliptic torus of G isomorphic to $\mathrm{Res}_{E/F}\mathbb{G}_m$. Its dual torus $\hat{T}$ is $\mathrm{Ind}_{E/F}(\mathbb{C}^\times)$, which is isomorphic to $(\mathbb{C}^\times)^n$ as a group. It is equipped with the induced action of the Weil group $\mathcal{W}_F$, which factors through the action of the Galois group Γ_F. We define the L-torus ${}^LT := \hat{T} \rtimes \mathcal{W}_F$ as the L-group of T.

We assume that the dual torus $\hat{T}$ is embedded into the L-group ${}^LG = \hat{G} \times \mathcal{W}_F$ of G, where $\hat{G} = \mathrm{GL}_n(\mathbb{C})$, with image $\mathcal{T}$. For convenience, we simply denote the image of $t \in \hat{T}$ by the embedding $\hat{T} \to \mathcal{T} \subset \hat{G}$ also by $t \in \mathcal{T}$. This embedding should be defined using the chosen splittings of G and $\hat{G}$. As we do not need the full detail of the definition of this embedding, we only refer to [LS87, Section 2.5] for details (or, when $(G,T) = (\mathrm{GL}_n, \mathrm{Res}_{E/F}\mathbb{G}_m)$, see [Tam, Section 6.1]). All we need to know is that we can always assume that the image $\mathcal{T}$ is the diagonal subgroup of $\hat{G}$.

With the embedding $\hat{T} \to \mathcal{T}$ chosen, an *admissible embedding* from ${}^LT$ to ${}^LG$ is a morphism of groups $I : {}^LT \to {}^LG$ of the form

$$I(t \rtimes w) = tI(1 \rtimes w) \text{ for all } t \rtimes w \in {}^LT.$$

Note that an admissible embedding maps $\mathcal{W}_F$ into $N_{\hat{G}}(\mathcal{T})$, i.e., the factor $I(1 \rtimes w)$ above lies in $N_{\hat{G}}(\mathcal{T})$. Two admissible embeddings I_1, I_2 are called $\mathrm{Int}(\mathcal{T})$-equivalent if there is $t \in \mathcal{T}$ such that

$$I_1(w) = tI_2(w)t^{-1} \text{ for all } w \in \mathcal{W}_F.$$

By [LS87, Section 2.6], admissible embeddings exist, and the collection of these embeddings can be described as follows.

Proposition 5.1. *The set of admissible embeddings from ${}^LT$ to ${}^LG$ is a $Z^1(\mathcal{W}_F, \hat{T})$-torsor, and the set of the* $\mathrm{Int}(\mathcal{T})$*-equivalence classes of these embeddings is an $H^1(\mathcal{W}_F, \hat{T})$-torsor.*

The idea in [LS87, Section 2.5] of constructing an admissible embedding is to choose a set of characters

$$\{\chi_\lambda\}_{\lambda\in\Phi}, \text{ where } \chi_\lambda : E_\lambda^\times \to \mathbb{C}^\times,$$

called χ-*data*, such that the following conditions hold.

Definition 5.2. (i) For each $\lambda \in \Phi$, we have $\chi_{-\lambda} = \chi_\lambda^{-1}$ and $\chi_{^w\lambda} = {}^w\chi_\lambda$ for all $w \in \mathcal{W}_F$.

(ii) If λ is symmetric, then $\chi|_{E_{\pm\lambda}^\times}$ equals the quadratic character $\delta_{E_\lambda/E_{\pm\lambda}}$ attached to the extension $E_\lambda/E_{\pm\lambda}$.

Remember that, in Section 2.1, we choose a subset $\mathcal{R}_\pm = \mathcal{R}_{\mathrm{sym}} \sqcup \mathcal{R}_{\mathrm{asym}/\pm}$ of Φ representing the orbits $\mathcal{W}_F\backslash\Phi_{\mathrm{sym}}$ and $\mathcal{W}_F\backslash\Phi_{\mathrm{asym}/\pm}$. Hence, by condition (i), the set of χ-data depends completely on the subset $\{\chi_\lambda\}_{\lambda\in\mathcal{R}_\pm}$. We still call such a subset a set of χ-data. Moreover, using Artin reciprocity [Tat79], we may regard each χ_λ as a character of the Weil group $\mathcal{W}_{E_\lambda}$.
Following the recipe in [LS87, Section 2.5], we can define an admissible embedding

$$I_{\{\chi_\lambda\}} : {}^LT \to {}^LG$$

depending on a given set of χ-data. In our present situation, we can describe the admissible embedding $I_{\{\chi_\lambda\}}$ in Proposition 5.3 below. We first recall the Langlands correspondence for the torus $T = \mathrm{Res}_{E/F}\mathbb{G}_m$, which is a bijection

$$\mathrm{Hom}(T(F), \mathbb{C}^\times) \to H^1(\mathcal{W}_F, \hat{T}). \tag{5.1}$$

Given a character ξ of $T(F) = E^\times$, we denote by $\tilde{\xi}$ a 1-cocycle in $Z^1(\mathcal{W}_F, \hat{T})$ whose class is the image of ξ under (5.1). Given χ-data $\{\chi_\lambda\}_{\lambda\in\mathcal{R}_\pm}$, we define

$$\mu := \mu_{\{\chi_\lambda\}} = \prod_{\lambda\in\mathcal{R}} \mathrm{Res}_{E^\times}^{E_\lambda^\times} \chi_\lambda,$$

where $\mathcal{R} = \mathcal{R}_\pm \sqcup (-\mathcal{R}_{\mathrm{asym}/\pm})$ is a subset representing $\Gamma_F\backslash\Phi$. It is easy to check that the product of the restricted characters is independent of representatives in $\mathcal{R}$, so we usually write

$$\mu = \prod_{[\lambda]\in\Gamma_F\backslash\Phi} \mathrm{Res}_{E^\times}^{E_\lambda^\times} \chi_\lambda.$$

Proposition 5.3 ([Tam, Proposition 6.5]). *For every character ξ of $E^\times$, the composition*

$$I_{\{\chi_\lambda\}} \circ \tilde{\xi} : \mathcal{W}_F \to {}^LT \to {}^LG \xrightarrow{proj} \mathrm{GL}_n(\mathbb{C})$$

is isomorphic to $\mathrm{Ind}_{E/F}(\xi \cdot \mu_{\{\chi_\lambda\}})$ *as a representation of* $\mathcal{W}_F$. □

We now define a analogous set of characters

$$\{\zeta_\lambda\}_{\lambda\in\Phi}, \text{ where } \zeta_\lambda : E_\lambda^\times \to \mathbb{C}^\times,$$

called *ζ-data*, such that the following conditions hold.

Definition 5.4. (i) For each $\lambda \in \Phi$, we have $\zeta_{-\lambda} = \zeta_\lambda^{-1}$ and $\zeta_{^w\lambda} = {}^w\zeta_\lambda$ for all $w \in \mathcal{W}_F$.

(ii) If λ is symmetric, then $\zeta|_{E_{\pm\lambda}^\times}$ is trivial.

We can view a set of ζ-data as the difference of two sets of χ-data. Motivated from Propositions 5.1 and 5.3, the product character

$$\nu := \nu_{\{\zeta_\lambda\}} = \prod_{[\lambda]\in\Gamma_F\backslash\Phi} \mathrm{Res}_{E^\times}^{E_\lambda^\times} \zeta_\lambda.$$

can be viewed as measuring the difference of two admissible embeddings. Recall that, similar to choosing $\mathcal{R}_\pm$, we can also choose $\mathcal{D}_\pm = \mathcal{D}_{\mathrm{sym}} \sqcup \mathcal{D}_{\mathrm{asym}/\pm}$ to be a subset of Γ_F/Γ_E consisting of representatives of $(\Gamma_E\backslash\Gamma_F/\Gamma_E)_{\mathrm{sym}}$ and $(\Gamma_E\backslash\Gamma_F/\Gamma_E)_{\mathrm{asym}/\pm}$ respectively, and obtain a bijection from Proposition 2.1,

$$\mathcal{R}_\pm = \mathcal{R}_{\mathrm{sym}} \bigsqcup \mathcal{R}_{\mathrm{asym}/\pm} \to \mathcal{D}_\pm = \mathcal{D}_{\mathrm{sym}} \bigsqcup \mathcal{D}_{\mathrm{asym}/\pm},\ \lambda = \left[\begin{smallmatrix} 1 \\ g \end{smallmatrix}\right] \mapsto g.$$

We usually denote by E_g and $E_{\pm g}$ the fields E_λ and $E_{\pm\lambda}$ respectively, if $g \in \mathcal{D}_\pm$ corresponds to $\lambda \in \mathcal{R}_\pm$. We also denote by ζ_g the character ζ_λ, and write

$$\nu := \nu_{\{\zeta_g\}} = \prod_{[g]\in(\mathcal{W}_E\backslash\mathcal{W}_F/\mathcal{W}_E)'} \mathrm{Res}_{E^\times}^{E_g^\times} \zeta_g.$$

5.2 Symmetric unramified zeta-data

We choose a specific ζ-data ζ_g for each $[g] = [\sigma^i\phi^{f/2}] \in (\mathcal{W}_E\backslash\mathcal{W}_F/\mathcal{W}_E)_{\mathrm{sym-unram}}$, base on the results from the χ-datum χ_g.
Notice that, since $E_g/E_{\pm g}$ is quadratic unramified, the norm group $N_{E_g/E_{\pm g}}(E_g^\times)$ has a decomposition

$$\boldsymbol{\mu}_{E_{\pm g}} \times \left\langle z_e^i z_{\phi^{f/2}} \varpi_E^2 \right\rangle \times U^1_{E_{\pm g}}$$

and we take a root of unity $z_0 \in \boldsymbol{\mu}_{E_g}$ such that

$$z_0\varpi_E \in E_{\pm g}^\times - N_{E_g/E_{\pm g}}(E_g^\times).$$

We only consider tamely ramified χ-data and ζ-data, i.e., we require that

$$\chi_g|_{U^1_{E_g}} \equiv 1 \text{ and } \zeta_g|_{U^1_{E_g}} \equiv 1.$$

Therefore, the Definition 5.2.(ii) of χ-data is explicitly (see [Tam, (7.6)])

$$\chi_g(\boldsymbol{\mu}_{E_{\pm g}}) = 1,\ \chi_g(z_e^i z_{\phi^{f/2}} \varpi_E^2) = 1,\ \text{and}\ \chi_g(z_0\varpi_E) = -1. \tag{5.2}$$

Hence, given a χ-datum χ_g, we can obtain a ζ-datum ζ_g easily by requiring

$$\zeta_g|_{\boldsymbol{\mu}_{E_g}} = \chi_g|_{\boldsymbol{\mu}_{E_g}} \text{ and } \zeta_g(\varpi_E) = -\chi_g(\varpi_E).$$

In [Tam, Section 7.4], in the cases when $\mathfrak{V}_{\xi,[g]} \cong \mathfrak{U}_{[g]}$ is non-trivial, we construct a χ-datum

$$\chi_g|_{\boldsymbol{\mu}_{E_g}} = t^1_{\boldsymbol{\mu}_g}(\mathfrak{U}_{[g]}) \text{ and } \chi_g(\varpi_E) = \begin{cases} -t_{\langle\varpi\rangle}(\mathfrak{U}_{[g]}) & \text{if } {}^{\sigma^i\phi^{f/2}}\varpi_E = \varpi_E, \\ t_{\langle\varpi\rangle}(\mathfrak{U}_{[g]}) & \text{otherwise.} \end{cases}$$

In other words, the character χ_g satisfies the conditions in (5.2). Hence the character

$$\zeta_g|_{\boldsymbol{\mu}_{E_g}} = t^1_{\boldsymbol{\mu}_g}(\mathfrak{U}_{[g]}) \text{ and } \zeta_g(\varpi_E) = \begin{cases} t_{\langle\varpi\rangle}(\mathfrak{U}_{[g]}) & \text{if } {}^{\sigma^i\phi^{f/2}}\varpi_E = \varpi_E, \\ -t_{\langle\varpi\rangle}(\mathfrak{U}_{[g]}) & \text{otherwise} \end{cases}$$

is a ζ-datum. This ζ-datum will be used in the next section.

5.3 Zeta-data associated to admissible characters

Given an admissible character ξ of $E^\times$ over F, we first assign, for each $[g] \in (\mathcal{W}_E\backslash\mathcal{W}_F/\mathcal{W}_E)_{\mathrm{asym}/\pm}$, the values of the ζ-data to

$$\zeta_{g,\xi}|_{\boldsymbol{\mu}_{E_g}} = \mathrm{sgn}_{\boldsymbol{\mu}_{E_g}}({}_A\mathfrak{V}_{[g]})\mathrm{sgn}_{\boldsymbol{\mu}_{E_g}}({}_{A^*}\mathfrak{V}_{[g]}) = t^1_{\boldsymbol{\mu}_g}({}_A\boldsymbol{\mathfrak{V}}_{\xi,[g]})t^1_{\boldsymbol{\mu}_g}({}_{A^*}\boldsymbol{\mathfrak{V}}_{\xi,[g]}).$$

In this way, the product of the characters

$$\zeta_{g,\xi}\zeta_{g^{-1},\xi} = \zeta_{g,\xi}\left(\zeta^g_{g,\xi}\right)^{-1} = \zeta_{g,\xi} \circ \left[\begin{smallmatrix} 1 & \\ & g \end{smallmatrix}\right]$$

has values

$$\begin{aligned}\left(\zeta_{g,\xi} \circ \left[\begin{smallmatrix} 1 & \\ & g \end{smallmatrix}\right]\right)|_{\boldsymbol{\mu}_E}(z) &= \mathrm{sgn}_{\boldsymbol{\mu}_E}({}_A\mathfrak{V}_{\xi,[g]})(\left[\begin{smallmatrix} 1 & \\ & g \end{smallmatrix}\right](z))\mathrm{sgn}_{\boldsymbol{\mu}_E}({}_{A^*}\mathfrak{V}_{\xi,[g]})(\left[\begin{smallmatrix} 1 & \\ & g \end{smallmatrix}\right](z)) \\ &= t^1_{\boldsymbol{\mu}}({}_A\boldsymbol{\mathfrak{V}}_{\xi,[g]})t^1_{\boldsymbol{\mu}}({}_{A^*}\boldsymbol{\mathfrak{V}}_{\xi,[g]})\end{aligned}$$

and

$$\begin{aligned}\left(\zeta_{g,\xi} \circ \left[\begin{smallmatrix} 1 & \\ & g \end{smallmatrix}\right]\right)(\varpi_E) &= \zeta_{g,\xi}|_{\boldsymbol{\mu}_E}\left(\left[\begin{smallmatrix} 1 & \\ & g \end{smallmatrix}\right](\varpi_E)\right) \\ &= \mathrm{sgn}_{\left[\begin{smallmatrix} 1 & \\ & g \end{smallmatrix}\right](\varpi_E)}({}_A\mathfrak{V}_{\xi,[g]})\mathrm{sgn}_{\left[\begin{smallmatrix} 1 & \\ & g \end{smallmatrix}\right](\varpi_E)}({}_{A^*}\mathfrak{V}_{\xi,[g]}) \\ &= t^1_{\langle\varpi\rangle}({}_A\boldsymbol{\mathfrak{V}}_{\xi,[g]})(\varpi_E)t^1_{\langle\varpi\rangle}({}_{A^*}\boldsymbol{\mathfrak{V}}_{\xi,[g]})(\varpi_E) \\ &= t_{\langle\varpi\rangle}({}_A\boldsymbol{\mathfrak{V}}_{\xi,[g]})t_{\langle\varpi\rangle}({}_{A^*}\boldsymbol{\mathfrak{V}}_{\xi,[g]}).\end{aligned} \tag{5.3}$$

We then assign, for each $[g] \in (\mathcal{W}_E\backslash\mathcal{W}_F/\mathcal{W}_E)_{\mathrm{asym}/\pm}$, arbitrary values to $\zeta_{g,\xi}(\varpi_E)$ and $\zeta_{g^{-1},\xi}(\varpi_E)$, as long as the product satisfies (5.3). (This phenomenon is comparable to [LS87, Lemma 3.3.A], as explained in [Tam, Remark 7.2].) It is routine to check that each $\zeta_{g,\xi}$ is a ζ-datum. Indeed, this checking is exactly the same as that in the χ-data case (see [Tam, Section 7.2]), since Definition 5.2.(i) is the same as that of χ-data.
We then assign values to the ζ-data for each $[g] \in (\mathcal{W}_E\backslash\mathcal{W}_F/\mathcal{W}_E)_{\mathrm{sym}}$ case by case.

5.3.1 Case when f is odd

Recall from (4.12) that

$$t^i_{\mathbf{\Gamma}}({}_A\mathfrak{V}_{\xi,\mathrm{sym}}) = t^i_{\mathbf{\Gamma}}({}_{A^*}\mathfrak{V}_{\xi,\mathrm{sym}}),\text{ for } i=0,1 \text{ and } \mathbf{\Gamma}=\boldsymbol{\mu},\langle\varpi\rangle. \tag{5.4}$$

We assign the ζ-data to the following values. If e is odd (so that m is odd since m divides e), we assign all $\zeta_{g,\xi}$ to be trivial. If e is even, then we just take all $\zeta_{g,\xi}$, $[g]\neq[\sigma^{e/2}]$, to be trivial and

$$\zeta_{\sigma^{e/2},\xi}|_{\boldsymbol{\mu}_E}\equiv 1 \text{ and } \zeta_{\sigma^{e/2},\xi}(\varpi_E)=(-1)^m.$$

To show that $\zeta_{\sigma^{e/2},\xi}$ is a ζ-datum, notice that since $N_{E/E_{\pm\sigma^{e/2}}}(\varpi_E) = -\varpi_{E_{\pm\sigma^{e/2}}} = -\varpi_E^2$, and since

$$\zeta_{\sigma^{e/2},\xi}(\varpi_E)^2=\zeta_{\sigma^{e/2},\xi}(-1)\zeta_{\sigma^{e/2},\xi}(-\varpi_E^2)=(1)(1)=1,$$

we can assign $\chi_{\sigma^{e/2},\xi}(\varpi_E)$ to either 1 or -1 to obtain a ζ-datum. By (5.4), we can rewrite our assigned ζ-data as

$$\begin{aligned}
\zeta_{g,\xi}|_{\boldsymbol{\mu}_E} &= t^1_{\boldsymbol{\mu}}({}_A\mathfrak{V}_{\xi,[g]})t^1_{\boldsymbol{\mu}}({}_{A^*}\mathfrak{V}_{\xi,[g]})\\
\zeta_{g,\xi}(\varpi_E) &= \begin{cases} t_{\langle\varpi\rangle}({}_A\mathfrak{V}_{\xi,[g]})t_{\langle\varpi\rangle}({}_{A^*}\mathfrak{V}_{\xi,[g]}) & \text{if } g\neq\sigma^{e/2},\\ (-1)^m t_{\langle\varpi\rangle}({}_A\mathfrak{V}_{\xi,[g]})t_{\langle\varpi\rangle}({}_{A^*}\mathfrak{V}_{\xi,[g]}) & \text{if } g=\sigma^{e/2}.\end{cases}
\end{aligned}\tag{5.5}$$

Note that $t_{\langle\varpi\rangle}({}_A\mathfrak{V}_{\xi,[\sigma^{e/2}]})$ is not defined (see the paragraph containing Formula (4.14)). In fact, we just take

$$t_{\langle\varpi\rangle}({}_A\mathfrak{V}_{\xi,[\sigma^{e/2}]})=t_{\langle\varpi\rangle}({}_{A^*}\mathfrak{V}_{\xi,[\sigma^{e/2}]})=1,$$

since ${}_A\mathfrak{V}_{\xi,[\sigma^{e/2}]}\cong{}_{A^*}\mathfrak{V}_{\xi,[\sigma^{e/2}]}$ by Proposition 3.2, and it is shown in [Tam, Proposition 5.3] that ${}_{A^*}\mathfrak{V}_{\xi,[\sigma^{e/2}]}$ is always trivial.
The product of ζ-data is equal to

$$\prod_{[g]\in(\mathcal{W}_E\backslash\mathcal{W}_F/\mathcal{W}_E)'}\zeta_{g,\xi}(\varpi_E)\equiv 1$$

$$\text{and}\quad \prod_{[g]\in(\mathcal{W}_E\backslash\mathcal{W}_F/\mathcal{W}_E)'}\zeta_{g,\xi}(\varpi_E)=\begin{cases}(-1)^m & \text{if } e \text{ is even},\\ 1 & \text{if } e \text{ is odd},\end{cases}$$

which is rewritten as

$$\prod_{[g]\in(\mathcal{W}_E\backslash\mathcal{W}_F/\mathcal{W}_E)'}\zeta_{g,\xi}|_{\boldsymbol{\mu}_E}\equiv t^1_{\boldsymbol{\mu}}({}_A\mathfrak{V}_\xi)t^1_{\boldsymbol{\mu}}({}_{A^*}\mathfrak{V}_\xi)$$

and

$$\prod_{[g]\in(\mathcal{W}_E\backslash\mathcal{W}_F/\mathcal{W}_E)'}\zeta_{g,\xi}(\varpi_E)=\begin{cases}(-1)^m t_{\langle\varpi\rangle}({}_A\mathfrak{V}_\xi)t_{\langle\varpi\rangle}({}_{A^*}\mathfrak{V}_\xi) & \text{if } e \text{ is even,}\\ t_{\langle\varpi\rangle}({}_A\mathfrak{V}_\xi)t_{\langle\varpi\rangle}({}_{A^*}\mathfrak{V}_\xi) & \text{if } e \text{ is odd.}\end{cases}$$

The product is equal to the rectifier given in (3.21),

$${}_D\nu_\xi|_{\boldsymbol{\mu}_E}\equiv 1 \text{ and } {}_D\nu_\xi(\varpi_E)=(-1)^{m(d-1)}=(-1)^{e-m},$$

when E/F is totally ramified.

5.3.2 Case when f is even

Let K be the maximal unramified extension of E/F. If we define ${}_{D_K}\nu_\xi$ to be the ramified part of ${}_D\nu_\xi$, which is also the rectifier corresponding to the admissible pair $(E/K,\xi)$, then we have

$${}_{D_K}\nu_\xi=\prod_{[g]\in(\mathcal{W}_E\backslash\mathcal{W}_K/\mathcal{W}_E)'}\zeta_{g,\xi}|_{E^\times},$$

as when f is odd, and in particular

$${}_{D_K}\nu_\xi(\varpi_E)=(-1)^{e-m_K},$$

where $m_K=\gcd(e,m)$. Therefore, our plan is to distribute the sign

$$(-1)^{e-m_K}(-1)^{n-m+f_\varpi-m_\varpi}=\begin{cases}1 & \text{if } m \text{ is even,}\\ (-1)^{e+f_\varpi+1} & \text{if } m \text{ is odd,}\end{cases}\tag{5.6}$$

to each $\zeta_{g,\xi}(\varpi_E)$, where $[g]$ is symmetric unramified, multiplying the product of t-factors $t_{\langle\varpi\rangle}({}_A\mathfrak{V}_{\xi,[g]})t_{\langle\varpi\rangle}({}_{A^*}\mathfrak{V}_{\xi,[g]})$. As before, we separate the cases according to the parity of m.

When m is even, recall from Proposition 4.9 that we have either

both ${}_A\mathfrak{V}_{\xi,k,\mathrm{sym}}$ and ${}_{A^*}\mathfrak{V}_{\xi,k,\mathrm{sym}}$ are trivial,

or

both ${}_A\mathfrak{V}_{\xi,k,\mathrm{sym}}$ and ${}_{A^*}\mathfrak{V}_{\xi,k,\mathrm{sym}}$ are isomorphic to $\mathfrak{U}_{k,\mathrm{sym}}$.

We assign the trivial ζ-data for all $[g]\in(\mathcal{W}_E\backslash\mathcal{W}_F/\mathcal{W}_E)_{\mathrm{sym-unram}}$, so that

$$\zeta_{g,\xi}|_{\boldsymbol{\mu}_E}=t^1_{\boldsymbol{\mu}}({}_A\mathfrak{V}_{\xi,[g]})t^1_{\boldsymbol{\mu}}({}_{A^*}\mathfrak{V}_{\xi,[g]}) \text{ and } \zeta_{g,\xi}(\varpi_E)=t_{\langle\varpi\rangle}({}_A\mathfrak{V}_{\xi,[g]})t_{\langle\varpi\rangle}({}_{A^*}\mathfrak{V}_{\xi,[g]}).$$

The product of $\zeta_{g,\xi}(\varpi_E)$ is trivial, or we can write

$$\prod_{[g]\in(\mathcal{W}_E\backslash\mathcal{W}_F/\mathcal{W}_E)'}\zeta_{g,\xi}|_{\boldsymbol{\mu}_E}=t^1_{\boldsymbol{\mu}}({}_A\mathfrak{V}_\xi)t^1_{\boldsymbol{\mu}}({}_{A^*}\mathfrak{V}_\xi)$$
$$\text{and}\quad\prod_{[g]\in(\mathcal{W}_E\backslash\mathcal{W}_F/\mathcal{W}_E)'}\zeta_{g,\xi}(\varpi_E)=t_{\langle\varpi\rangle}({}_A\mathfrak{V}_\xi)t_{\langle\varpi\rangle}({}_{A^*}\mathfrak{V}_\xi).$$

Note that in the second product, the sign without t-factors is equal to (5.6), which is just 1.
When m is odd, we have

$${}_A\mathfrak{V}_{\xi,\mathrm{sym-unram}}\oplus{}_{A^*}\mathfrak{V}_{\xi,\mathrm{sym-unram}}=\mathfrak{U}_{\mathrm{sym-unram}}.$$

We then assign the ζ-data to be

$$\zeta_{g,\xi}|_{\boldsymbol{\mu}_E}=t^1_{\boldsymbol{\mu}}({}_A\mathfrak{V}_{\xi,[g]})t^1_{\boldsymbol{\mu}}({}_{A^*}\mathfrak{V}_{\xi,[g]})=t^1_{\boldsymbol{\mu}}(\mathfrak{U}_{[g]})$$
$$\text{and }\zeta_{g,\xi}(\varpi_E)=-t_{\langle\varpi\rangle}({}_A\mathfrak{V}_{\xi,[g]})t_{\langle\varpi\rangle}({}_{A^*}\mathfrak{V}_{\xi,[g]})=-t_{\langle\varpi\rangle}(\mathfrak{V}_{[g]}).$$

for all symmetric unramified $[g]$ except the one which stabilizes ϖ_E, in which we assign

$$\zeta_{g,\xi}|_{\boldsymbol{\mu}_E}=t^1_{\boldsymbol{\mu}}({}_A\mathfrak{V}_{\xi,[g]})t^1_{\boldsymbol{\mu}}({}_{A^*}\mathfrak{V}_{\xi,[g]})=t^1_{\boldsymbol{\mu}}(\mathfrak{U}_{[g]})$$
$$\text{and }\zeta_{g,\xi}(\varpi_E)=t_{\langle\varpi\rangle}({}_A\mathfrak{V}_{\xi,[g]})t_{\langle\varpi\rangle}({}_{A^*}\mathfrak{V}_{\xi,[g]})=t_{\langle\varpi\rangle}(\mathfrak{V}_{[g]}).$$

In Section 5.2, we checked that the above characters give rise to ζ-data. The product of ζ is hence

$$\prod_{[g]\in(\mathcal{W}_E\backslash\mathcal{W}_F/\mathcal{W}_E)'}\zeta_{g,\xi}|_{\boldsymbol{\mu}_E}=t^1_{\boldsymbol{\mu}}({}_A\mathfrak{V}_\xi)t^1_{\boldsymbol{\mu}}({}_{A^*}\mathfrak{V}_\xi)$$
$$\text{and}\quad\prod_{[g]\in(\mathcal{W}_E\backslash\mathcal{W}_F/\mathcal{W}_E)'}\zeta_{g,\xi}(\varpi_E)=(-1)^{e+f_\varpi+1}t_{\langle\varpi\rangle}({}_A\mathfrak{V}_\xi)t_{\langle\varpi\rangle}({}_{A^*}\mathfrak{V}_\xi),$$

by Proposition 2.6. Again in the second product, the sign without t-factors is equal to (5.6).

5.4 The main theorem

To summarize, we verified the following theorem.

Theorem 5.5. *Let ξ be an admissible character of $E^\times$ over F.*

(i) Let ${}_A\mathfrak{V}_\xi$ (resp. ${}_{A^*}\mathfrak{V}_\xi$) be the finite symplectic module defined by ξ when $G(F)=\mathrm{GL}_m(D)$ (resp. when $G^*(F)=\mathrm{GL}_n(F)$). The following conditions define a collection of ζ-data $\{\zeta_{g,\xi}\}_{g\in\mathcal{D}_{\mathrm{asym}/\pm}\sqcup\mathcal{D}_{\mathrm{sym}}}$.

(a) All $\zeta_{g,\xi}$ are tamely ramified.

(b) If $g \in \mathcal{D}_{\mathrm{asym}/\pm}$, then

$$\zeta_{g,\xi}|_{\boldsymbol{\mu}_{E_g}} = t^1_{\boldsymbol{\mu}_g}({}_A\mathfrak{V}_{\xi,[g]})t^1_{\boldsymbol{\mu}_g}({}_{A^*}\mathfrak{V}_{\xi,[g]})$$

and

$$\begin{aligned}&\zeta_{g,\xi}(\varpi_E) \text{ can be any value satisfying}\\ &\zeta_{g,\xi}(\varpi_E)\zeta_{g^{-1},\xi}(\varpi_E) = t_{\langle\varpi\rangle}({}_A\mathfrak{V}_{\xi,[g]})t_{\langle\varpi\rangle}({}_{A^*}\mathfrak{V}_{\xi,[g]}).\end{aligned}$$

(c) If $g \in \mathcal{D}_{\mathrm{sym}}$, then

$$\zeta_{g,\xi}|_{\boldsymbol{\mu}_{E_g}} = t^1_{\boldsymbol{\mu}_g}({}_A\mathfrak{V}_{\xi,[g]})t^1_{\boldsymbol{\mu}_g}({}_{A^*}\mathfrak{V}_{\xi,[g]})$$

and

$$\zeta_{g,\xi}(\varpi_E) = \epsilon_g t_{\langle\varpi\rangle}({}_A\mathfrak{V}_{\xi,[g]})t_{\langle\varpi\rangle}({}_{A^*}\mathfrak{V}_{\xi,[g]}),$$

where ϵ_g is equal to 1 if $g \in (\mathcal{D}_{\mathrm{sym-ram}} - \{\sigma^{e/2}\}) \cup \mathcal{W}_{F[\varpi_E]}$ and is equal to $(-1)^m$ if $g \in (\mathcal{D}_{\mathrm{sym-unram}} - \mathcal{W}_{F[\varpi_E]}) \cup \{\sigma^{e/2}\}$.

(ii) Let ${}_D\nu_\xi$ be the rectifier of ξ and $\{\zeta_{g,\xi}\}_{g\in\mathcal{D}_\pm}$ be the ζ-data in (i), then

$${}_D\nu_\xi = \prod_{[g]\in(\mathcal{W}_E\backslash\mathcal{W}_F/\mathcal{W}_E)'} \zeta_{g,\xi}|_{E^\times}.$$

□

Remark 5.6. As long as the F-dimension of the division algebra D is fixed, the rectifier ${}_D\nu_\xi$ is independent of the Hasse-invariant $h = h(D)$ of D, as stated in [BH11, Theorem C]. This is because the modules ${}_A\mathfrak{V}_\xi$, where $A = \mathrm{Mat}_n(D)$ and D ranges over all division algebra with same F-dimension, are all isomorphic to each other. Similarly, the ζ-data $\{\zeta_{g,\xi}\}$ are independent of $h(D)$. □

5.5 Functorial property

Let K be an intermediate subfield in E/F, and write

$$n_K = n(E/K) = f_K e_K = f(E/K)e(E/K) \text{ and } m_K = \gcd(m, n_K).$$

Similar to Section 3.2, we have the Jacquet-Langlands correspondence

$$JL_K : \mathcal{A}^{\mathrm{et}}_{n_K}(K) \to \mathcal{A}^{\mathrm{et}}_{m_K}(D_K),$$

between essentially tame supercuspidal representations of $G(F)_K = \mathrm{GL}_{m_K}(D_K)$ and its split inner form $G^*(F)_K = \mathrm{GL}_{n_K}(K)$. We can parametrize both collections by the admissible pairs in $P_{n_K}(K)$, and obtain the rectifier map

$${}_{D_K}\nu : P_{n_K}(K) \xrightarrow{{}_K\Pi} \mathcal{A}^{\mathrm{et}}_{n_K}(K) \xrightarrow{JL_K} \mathcal{A}^{\mathrm{et}}_{m_K}(D_K) \xrightarrow{{}_{D_K}\Pi^{-1}} P_{n_K}(K),$$

such that

$$_{D_K}\nu(E/K,\xi) = (E/F, \xi \cdot {}_{D_K}\nu_\xi).$$

for a tamely ramified character $_{D_K}\nu_\xi$ of $E^\times$ for each pair $(E/K,\xi) \in P_{n_K}(K)$. With the embedding condition for $(E_0K, \mathfrak{A}_K)$ as discussed in Section 2.5, we define the subgroups (see also [BH11, 3.2 Proposition])

$$H^1_K = H^1(\Xi, \mathfrak{A}) \cap G(F)_K \text{ and } J^1_K = J^1(\Xi, \mathfrak{A}) \cap G(F)_K$$

Each subgroup above admits a similar factorization as in (3.10). We then obtain

$$_{A_K}\mathfrak{V}_\xi = J^1_K/H^1_K = {}_A\mathfrak{V}_\xi \cap \mathfrak{U}_{E/K}$$

and similarly for $_{A^*_K}\mathfrak{V}_\xi$.

Denote $\Psi_{E/K} = E^\times/K^\times U^1_E$, and view $_{A_K}\mathfrak{V}_\xi$ and $_{A^*_K}\mathfrak{V}_\xi$ as $\mathbf{k}_K\Psi_{E/K}$-submodules of $\mathfrak{U}_{E/K}$. Denote the subgroups of $\Psi_{E/K}$ by

$$\begin{aligned} &\boldsymbol{\mu}_{E/K} = \boldsymbol{\mu}_E/\boldsymbol{\mu}_K \text{ and} \\ &\langle\varpi\rangle_{E/K} = \text{the subgroup generated by the image of } \varpi_E. \end{aligned}$$

Using the results in Section 3.8, with the base field changed from F to K, the values of $_{D_K}\nu_\xi$ is given by

$$\begin{aligned} {}_{D_K}\nu_\xi|_{\boldsymbol{\mu}_E} &= t^1_{\boldsymbol{\mu}_{E/K}}({}_{A_K}\mathfrak{V}_\xi)t^1_{\boldsymbol{\mu}_{E/K}}({}_{A^*_K}\mathfrak{V}_\xi) \\ \text{and } {}_{D_K}\nu_\xi(\varpi_E) &= (-1)^{n_K - m_K + f_{\varpi,K} - m_{\varpi,K}} t_{\langle\varpi\rangle_{E/K}}({}_{A_K}\mathfrak{V}_\xi)t_{\langle\varpi\rangle_{E/K}}({}_{A^*_K}\mathfrak{V}_\xi) \end{aligned}$$

for a prime element $\varpi_E \in E_0K$ (see the beginning of Section 3.8). Here

$$\begin{aligned} &f_{\varpi,K} = f(E/K[\varpi_E]) \text{ and} \\ &m_{\varpi,K} = \gcd(m_K, f_{\varpi,K}) = \gcd(m, n_K, f(E/K[\varpi_E])). \end{aligned}$$

Now suppose that $(E/F,\xi) \in P_n(F)$. By the definition of admissibility, we can regard ξ as an admissible character over K and form the pair $(E/K,\xi) \in P_{n_K}(K)$.

Proposition 5.7. *In this situation, we have*

$$_{D_K}\nu_\xi = \prod_{[g] \in (\mathcal{W}_E \backslash \mathcal{W}_K/\mathcal{W}_E)'} \zeta_{g,\xi}|_{E^\times}.$$

Proof. Notice that if $\mathfrak{V}$ is a $\mathbf{k}_F\Psi_{E/F}$-submodule of $\mathfrak{U}_{E/K}$, we have

$$t^1_{\boldsymbol{\mu}_{E/K}}(\mathfrak{V}) = t^1_{\boldsymbol{\mu}_{E/F}}(\mathfrak{V}) \text{ and } t_{\langle\varpi\rangle_{E/K}}(\mathfrak{V}) = t_{\langle\varpi\rangle_{E/F}}(\mathfrak{V}),$$

where $\boldsymbol{\mu}_{E/F}$ and $\langle\varpi\rangle_{E/F}$ are just $\boldsymbol{\mu}$ and $\langle\varpi\rangle$ respectively considered in (3.20). Hence we have

$$_{D_K}\nu_\xi|_{\boldsymbol{\mu}_E} = \prod_{[g] \in (\mathcal{W}_E \backslash \mathcal{W}_K/\mathcal{W}_E)'} \zeta_{g,\xi}|_{\boldsymbol{\mu}_E}.$$

It remains to consider the values of both characters at ϖ_E. Notice that the left side has value

$$(-1)^{n_K-m_K+f_{\varpi,K}-m_{\varpi,K}}t_{\langle\varpi\rangle_{E/K}}(_{A_K}\mathfrak{V}_\xi)t_{\langle\varpi\rangle_{E/K}}(_{A_K^*}\mathfrak{V}_\xi),$$

while the right side has value

$$(\text{a sign})\cdot\prod_{[g]\in(\mathcal{W}_E\backslash\mathcal{W}_K/\mathcal{W}_E)'}t_{\langle\varpi\rangle}(_{A_K}\mathfrak{V}_{\xi,[g]})t_{\langle\varpi\rangle}(_{A_K^*}\mathfrak{V}_{\xi,[g]}).$$

The t-factors on both sides are clearly equal. We will recall, by Theorem 5.5, the values of the sign on the right side in different cases and show that, in each case, this sign is equal to the one on the left side.

We first consider when f_K is odd, which can be reduced to the case when E/K is totally ramified. We further separate into cases.

- When e is odd, or when e is even and e_K is odd, then m_K is also odd. The sign on the left is $(-1)^{e_K-m_K}=1$, while that on the right is also 1 since $\sigma^{e/2}\notin\mathcal{W}_K$.
- When e_K is even, the sign on the left is $(-1)^{e_K-m_K}=(-1)^m$ since $m_K\equiv m \mod 2$, while that on the right is $(-1)^m$ since $\sigma^{e/2}\in\mathcal{W}_K$.

We then consider then f_K is even. Let L be the maximal unramified extension of E/K. We recall, after disregarding the symmetric ramified component (as we did at the beginning of Sub-section 5.3.2), the sign on the left is equal to (see (5.6))

$$(-1)^{e_K-m_L}(-1)^{n_K-m_K+f_{\varpi,K}-m_{\varpi,K}}=\begin{cases}1 & \text{if } m_K \text{ is even,}\\ (-1)^{e_K+f_{\varpi,K}+1} & \text{if } m_K \text{ is odd.}\end{cases}$$

Recall from Proposition 2.6 that the number $e_K+f_{\varpi,K}+1$ is just the cardinality of

$$(\Gamma_E\backslash\Gamma_K/\Gamma_E)_{\text{sym-unram}}-\Gamma_{K[\varpi_E]}.$$

Hence by Theorem 5.5, the sign on the right side is

$$(-1)^{m(e_K+f_{\varpi,K}+1)}.$$

By knowing that $m_K\equiv m \mod 2$, the sign above is equal to the one on the left side. $\square$

Acknowledgements

This work was started when the author was receiving funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC Grant agreement no.290766 (AAMOT) during his visit in IMJ, Paris. It was then finished when he became a post-doc in McMaster University. He would like to thank Vincent Sécherre for a few discussions in Paris, and Shaun Stevens for reading the manuscript and his comments. He would also like to thank the referee for the careful reading.

References

[Bad02] Alexandru Ioan Badulescu, *Correspondance de Jacquet-Langlands pour les corps locaux de caractéristique non nulle*, Ann. Sci. École Norm. Sup. (4) 35 (2002), no. 5, 695–747. MR 1951441 (2004i:11048)

[BF83] Colin J. Bushnell and Albrecht Fröhlich, *Gauss sums and p-adic division algebras*, Lecture Notes in Mathematics, vol. 987, Springer-Verlag, Berlin-New York, 1983. MR 701540 (84m:12017)

[BF85] C. J. Bushnell and A. Fröhlich, *Nonabelian congruence Gauss sums and p-adic simple algebras*, Proc. London Math. Soc. (3) 50 (1985), no. 2, 207–264. MR 772712 (86g:11071)

[BH96] Colin J. Bushnell and Guy Henniart, *Local tame lifting for* $\mathrm{GL}(N)$*. I. Simple characters*, Inst. Hautes Études Sci. Publ. Math. (1996), no. 83, 105–233. MR 1423022 (98m:11129)

[BH05a] ______, *The essentially tame local Langlands correspondence. I*, J. Amer. Math. Soc. 18 (2005), no. 3, 685–710 (electronic). MR 2138141 (2006a:22014)

[BH05b] ______, *The essentially tame local Langlands correspondence. II. Totally ramified representations*, Compos. Math. 141 (2005), no. 4, 979–1011. MR 2148193 (2006c:22017)

[BH10] ______, *The essentially tame local Langlands correspondence, III: the general case*, Proc. Lond. Math. Soc. (3) 101 (2010), no. 2, 497–553. MR 2679700

[BH11] ______, *The essentially tame Jacquet-Langlands correspondence for inner forms of* $\mathrm{GL}(n)$, Pure Appl. Math. Q. 7 (2011), no. 3, Special Issue: In honor of Jacques Tits, 469–538. MR 2848585 (2012f:22027)

[BHL10] Colin J. Bushnell, Guy Henniart, and Bertrand Lemaire, *Caractère et degré formel pour les formes intérieures de* $\mathrm{GL}(n)$ *sur un corps local de caractéristique non nulle*, Manuscripta Math. 131 (2010), no. 1-2, 11–24. MR 2574988 (2010m:22013)

[BK93] C.J. Bushnell and P.C. Kutzko, *The admissible dual of GL(N) via compact open subgroups*, Annals of mathematics studies, Princeton University Press, 1993.

[Bro96] Paul Broussous, *Extension du formalisme de Bushnell et Kutzko au cas d'une algèbre à division*, C. R. Acad. Sci. Paris Sér. I Math. 323 (1996), no. 3, 225–228. MR 1404764 (97f:22028)

[BSS12] P. Broussous, V. Sécherre, and S. Stevens, *Smooth representations of* $\mathrm{GL}_m(D)$ *V: Endo-classes*, Doc. Math. 17 (2012), 23–77. MR 2889743

[DKV84] P. Deligne, D. Kazhdan, and M.-F. Vignéras, *Représentations des algèbres centrales simples p-adiques*, Representations of reductive groups over a local field, Travaux en Cours, Hermann, Paris, 1984, pp. 33–117. MR 771672 (86h:11044)

[Gra07] Martin Grabitz, *Simple characters for principal orders in* $\mathrm{M}_m(D)$, J. Number Theory 126 (2007), no. 1, 1–51. MR 2347425 (2008k:11125)

[Gre55] J. A. Green, *The characters of the finite general linear groups*, Trans. Amer. Math. Soc. 80 (1955), 402–447. MR 0072878 (17,345e)

[How77] Roger E. Howe, *Tamely ramified supercuspidal representations of* Gl_n, Pacific J. Math. 73 (1977), no. 2, 437–460. MR 0492087 (58 #11241)

[JL70] H. Jacquet and R. P. Langlands, *Automorphic forms on* GL(2), Lecture Notes in Mathematics, Vol. 114, Springer-Verlag, Berlin, 1970. MR 0401654 (53 #5481)

[KS99] Robert E. Kottwitz and Diana Shelstad, *Foundations of twisted endoscopy*, Astérisque (1999), no. 255, vi+190. MR 1687096 (2000k:22024)

[Lan94] Serge Lang, *Algebraic number theory*, second ed., Graduate Texts in Mathematics, vol. 110, Springer-Verlag, New York, 1994. MR 1282723 (95f:11085)

[LS87] R. P. Langlands and D. Shelstad, *On the definition of transfer factors*, Math. Ann. 278 (1987), no. 1-4, 219–271. MR 909227 (89c:11172)

[Moy86] Allen Moy, *Local constants and the tame Langlands correspondence*, Amer. J. Math. 108 (1986), no. 4, 863–930. MR 853218 (88b:11081)

[NM43] Tadasi Nakayama and Yozô Matsushima, *Über die multiplikative Gruppe einer p-adischen Divisionsalgebra*, Proc. Imp. Acad. Tokyo 19 (1943), 622–628. MR 0014081 (7,238a)

[Rei03] I. Reiner, *Maximal orders*, London Mathematical Society Monographs. New Series, vol. 28, The Clarendon Press Oxford University Press, Oxford, 2003, Corrected reprint of the 1975 original, With a foreword by M. J. Taylor. MR 1972204 (2004c:16026)

[Rog83] Jonathan D. Rogawski, *Representations of* $\mathrm{GL}(n)$ *and division algebras over a p-adic field*, Duke Math. J. 50 (1983), no. 1, 161–196. MR 700135 (84j:12018)

[Séc04] Vincent Sécherre, *Représentations lisses de* $\mathrm{GL}(m, D)$. *I. Caractères simples*, Bull. Soc. Math. France 132 (2004), no. 3, 327–396. MR 2081220 (2005f:22027)

[Séc05a] ———, *Représentations lisses de* $\mathrm{GL}(m, D)$*. II.* β*-extensions*, Compos. Math. 141 (2005), no. 6, 1531–1550. MR 2188448 (2006j:22018)

[Séc05b] ———, *Représentations lisses de* $\mathrm{GL}_m(D)$*. III. Types simples*, Ann. Sci. École Norm. Sup. (4) 38 (2005), no. 6, 951–977. MR 2216835 (2007a:22010)

[SS08] V. Sécherre and S. Stevens, *Représentations lisses de* $GL(m, D)$*. IV. Représentations supercuspidales*, J. Inst. Math. Jussieu 7 (2008), no. 3, 527–574. MR 2427423 (2009d:22023)

[Tam] Kam-Fai Tam, *Admissible embeddings of L-tori and the essentially tame local Langlands correspondence, to appear in* International Mathematics Research Notices.

[Tat79] J. Tate, *Number theoretic background*, Automorphic forms, representations and L-functions (Proc. Sympos. Pure Math., Oregon State Univ., Corvallis, Ore., 1977), Part 2, Proc. Sympos. Pure Math., XXXIII, Amer. Math. Soc., Providence, R.I., 1979, pp. 3–26. MR 546607 (80m:12009)

[Zin92] Ernst-Wilhelm Zink, *Representation theory of local division algebras*, J. Reine Angew. Math. 428 (1992), 1–44. MR 1166506 (93i:11141)

[Zin99] ———, *More on embeddings of local fields in simple algebras*, J. Number Theory 77 (1999), no. 1, 51–61. MR 1695699 (2001d:11116)

K.-F. Tam
Dept. of Math. & Stat.
McMaster University
1280 Main Street West
Hamilton, Ontario, Canada
L8S 4K1
geotam@math.mcmaster.ca

Remarks on L^p-Boundedness of Wave Operators for Schrödinger Operators with Threshold Singularities

K. Yajima [1]

Received: February 24, 2016

Communicated by Heinz Siedentop

Abstract. We consider the continuity property in Lebesgue spaces $L^p(\mathbb{R}^m)$ of the wave operators $W_\pm$ of scattering theory for Schrödinger operators $H = -\Delta + V$ on $\mathbb{R}^m$, $|V(x)| \leq C\langle x\rangle^{-\delta}$ for some $\delta > 2$ when H is of exceptional type, i.e. $\mathcal{N} = \{u \in \langle x\rangle^s L^2(\mathbb{R}^m)\colon (1 + (-\Delta)^{-1}V)u = 0\} \neq \{0\}$ for some $1/2 < s < \delta - 1/2$. It has recently been proved by Goldberg and Green for $m \geq 5$ that $W_\pm$ are in general bounded in $L^p(\mathbb{R}^m)$ for $1 \leq p < m/2$, for $1 \leq p < m$ if all $\phi \in \mathcal{N}$ satisfy $\int_{\mathbb{R}^m} V\phi dx = 0$ and, for $1 \leq p < \infty$ if $\int_{\mathbb{R}^m} x_i V\phi dx = 0$, $i = 1, \ldots, m$ in addition. We make the results for $p > m/2$ more precise and prove in particular that these conditions are also necessary for the stated properties of $W_\pm$. We also prove that, for $m = 3$, $W_\pm$ are bounded in $L^p(\mathbb{R}^3)$ for $1 < p < 3$ and that the same holds for $1 < p < \infty$ if and only if all $\phi \in \mathcal{N}$ satisfy $\int_{\mathbb{R}^3} V\phi dx = 0$ and $\int_{\mathbb{R}^3} x_i V\phi dx = 0$, $i = 1, 2, 3$, simultaneously.

2010 Mathematics Subject Classification: 35P25, 81U05, 47A40.

1 Introduction

Let $H_0 = -\Delta$ be the free Schrödinger operator on the Hilbert space $\mathcal{H} = L^2(\mathbb{R}^m)$ with domain $D(H_0) = \{u \in \mathcal{H}\colon -\Delta u \in \mathcal{H}\}$ and $H = H_0 + V$, V being the multiplication operator with the real measurable function $V(x)$ which satisfies

$$|V(x)| \leq C\langle x\rangle^{-\delta} \ \text{ for some } \ \delta > 2,\ \langle x\rangle = (1 + |x|^2)^{\frac{1}{2}}. \tag{1.1}$$

Then, H is selfadjoint in $\mathcal{H}$ with a core $C_0^\infty(\mathbb{R}^m)$ and it satisfies the following properties (see e.g. [18, 19, 21, 22, 23]):

[1] Supported by JSPS grant in aid for scientific research No. 22340029

(i) The spectrum $\sigma(H)$ of H consists of the absolutely continuous (AC for short) part $[0,\infty)$ and a finite number of non-positive eigenvalues of finite multiplicities.

We write $\mathcal{H}_{ac}(H)$ for the AC spectral subspace of $\mathcal{H}$ for H, H_{ac} for the part of H in $\mathcal{H}_{ac}(H)$ and $P_{ac}(H)$ for the orthogonal projection onto $\mathcal{H}_{ac}(H)$.

(ii) Wave operators $W_\pm = \lim_{t\to\pm\infty} e^{itH}e^{-itH_0}$ defined by strong limits exist and are complete, viz. $\operatorname{Image} W_\pm = \mathcal{H}_{ac}(H)$. They are unitary from $\mathcal{H}$ onto $\mathcal{H}_{ac}(H)$ and intertwine H_{ac} and H_0. Hence, for Borel functions f,

$$f(H)P_{ac}(H) = W_\pm f(H_0)W_\pm^*. \tag{1.2}$$

If follows that various mapping properties of $f(H)P_{ac}$ may be deduced from those of $f(H_0)$ if the corresponding ones of $W_\pm$ are known. In particular, if $W_\pm \in \mathbf{B}(L^p(\mathbb{R}^m))$ for $1 \le p_1 \le p \le p_2 < \infty$, then $W_\pm^* \in \mathbf{B}(L^q(\mathbb{R}^m))$ for $q_2 \le q \le q_1$, $1/p_j + 1/q_j = 1$, $j = 1, 2$, and

$$\|f(H)P_{ac}(H)\|_{\mathbf{B}(L^q,L^p)} \le C_{pq}\|f(H_0)\|_{\mathbf{B}(L^q,L^p)}, \tag{1.3}$$

for these p and q with C_{pq} which are *independent of* f. We define the Fourier and the conjugate Fourier transforms $\mathcal{F}u(\xi)$ and $\mathcal{F}^*u(\xi)$ respectively by

$$\mathcal{F}u(\xi) = \int_{\mathbb{R}^m} e^{-ix\xi}u(x)dx \text{ and } \mathcal{F}^*u(\xi) = \frac{1}{(2\pi)^m}\int_{\mathbb{R}^m} e^{ix\xi}u(x)dx.$$

We also write $\hat{u}(\xi)$ for $\mathcal{F}u(\xi)$.
The intertwining property (1.2) may be made more precise. Wave operators $W_\pm$ are transplantations ([24]) of the complete set of (generalized) eigenfunctions $\{e^{ix\xi} \colon \xi \in \mathbb{R}^m\}$ of $-\Delta$ by those of out-going and in-coming scattering eigenfunctions $\{\varphi_\pm(x,\xi) \colon \xi \in \mathbb{R}^m\}$ of $H = -\Delta + V$ ([19]):

$$W_\pm u(x) = \mathcal{F}_\pm^*\mathcal{F}u(x) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} \varphi_\pm(x,\xi)\hat{u}(\xi)d\xi,$$

where $\mathcal{F}_\pm$ and $\mathcal{F}_\pm^*$ are the generalized Fourier transforms associated with $\{\varphi_\pm(x,\xi) \colon \xi \in \mathbb{R}^m\}$ and the conjugate ones defined respectively by

$$\mathcal{F}_\pm u(\xi) = \int_{\mathbb{R}^d} \overline{\varphi_\pm(x,\xi)}u(x)dx, \quad \mathcal{F}_\pm^*u(\xi) = \frac{1}{(2\pi)^m}\int_{\mathbb{R}^d} \varphi_\pm(x,\xi)u(x)dx.$$

They satisfy $\mathcal{F}_\pm^*\mathcal{F}_\pm u = u$ for $u \in \mathcal{H}_{ac}(H)$ and, $\mathcal{F}_\pm\mathcal{F}_\pm^* u = u$ for $u \in L^2(\mathbb{R}^m)$. We define $F(D) \equiv \mathcal{F}^*M_F\mathcal{F}$ and $F(D_\pm) \equiv \mathcal{F}_\pm^*M_F\mathcal{F}_\pm$ for Borel functions F on $\mathbb{R}^m$ where M_F is the multiplication with $F(\xi)$. Then,

$$F(D_\pm) = W_\pm F(D)W_\pm^*u, \quad u \in \mathcal{H}_{ac}(H)$$

and $W_\pm$ transplant estimates for $F(D)$ in L^p-spaces to $F(D_\pm)$.

In this paper we are interested in the problem whether or not $W_\pm$ are bounded in $L^p(\mathbb{R}^m)$. This will almost automatically imply the same property in Sobolev spaces $W^{k,p}(\mathbb{R}^m) = \{u \in L^p(\mathbb{R}^m) \colon \partial^\alpha u \in L^p(\mathbb{R}^m)\}$ for integers $0 \le k \le 2$ (see Section 7 of [8]).

There is now a large literature on this problem ([3, 4, 6, 8, 27, 31, 15, 16, 29, 33]) and it is well known that the answer depends on the spectral properties of H at 0, the bottom of the AC spectrum of H. We define

$$\mathcal{E} = \{u \in H^2(\mathbb{R}^m) \colon (-\Delta + V)u = 0\}, \tag{1.4}$$

the eigenspace of H with eigenvalue 0 and, for $1/2 < s < \delta - 1/2$,

$$\mathcal{N} = \{u \in \langle x \rangle^s L^2(\mathbb{R}^m) : (1 + (-\Delta)^{-1}V)u = 0\} = 0. \tag{1.5}$$

Functions ϕ in $\mathcal{N}$ satisfy $-\Delta\phi + V\phi = 0$ for $x \in \mathbb{R}^m$. The space $\mathcal{N}$ is finite dimensional, independent of $1/2 < s < \delta - 1/2$, $\mathcal{E} \subset \mathcal{N}$ and, if $m \ge 5$, $\mathcal{E} = \mathcal{N}$ ([14]). The operator H is said be of *generic type* if $\mathcal{N} = \{0\}$ and of *exceptional type* otherwise. When H is of *generic type*, we have rather satisfactory results (though there is much space for improving conditions on V) and it has been proved that $W_\pm$ are bounded in $L^p(\mathbb{R}^m)$ for all $1 \le p \le \infty$ if $m \ge 3$ and, for all $1 < p < \infty$ if $m = 1$ and $m = 2$ under various smoothness and decay at infinity assumptions on V (see [4] for the best result when $m = 3$); but they are in general *not* bounded in $L^1(\mathbb{R}^1)$ or $L^\infty(\mathbb{R}^1)$ when $m = 1$ ([27]).

When H is of exceptional type, it is long known that the same results hold when $m = 1$ (see [27, 3, 6]). For higher dimensions $m \ge 3$, it is first shown ([33, 8]) that $W_\pm$ are bounded in $L^p(\mathbb{R}^m)$ for $3/2 < p < 3$ if $m = 3$ and for $\frac{m}{m-2} < p < \frac{m}{2}$ if $m \ge 5$, which is subsequently extended to $1 < p < 3$ for $m = 3$ and $1 < p < m/2$ for $m \ge 5$ ([34]). Then, recently, Goldberg and Green ([10]) have substantially improved these results by proving the following theorem for $m \ge 5$. In what follows in this paper, we assume $m \ge 3$ and V satisfies the following assumption. The constant m_* is defined by

$$m_* = (m-1)/(m-2).$$

Assumption 1.1. *V is a real valued measurable function such that*

(1) $\mathcal{F}(\langle x \rangle^{2\sigma} V) \in L^{m_*}$ *for some* $\sigma > 1/m_*$.

(2) $|V(x)| \le C\langle x \rangle^{-\delta}$ *for some* $\delta > \begin{cases} m+4, & \textit{if } 3 \le m \le 7, \\ m+3, & \textit{if } m \ge 8 \end{cases}$ *and* $C > 0$.

The condition (1) requires certain smoothness on V.

We write $\langle u, v \rangle = \int_{\mathbb{R}^m} \overline{u(x)} v(x) dx$ and define subspaces $\mathcal{E}_1 \subset \mathcal{E}_0 \subset \mathcal{N}$ respectively by

$$\mathcal{E}_0 = \{\phi \in \mathcal{N} \colon \langle V, \phi \rangle = 0\}, \quad \mathcal{E}_1 = \{\phi \in \mathcal{E}_0 \colon \langle xV, \phi \rangle = 0\}, \tag{1.6}$$

where $\langle xV, \phi \rangle = 0$ means $\langle x_i V, \phi \rangle = 0$ for all $1 \le i \le m$. We have $\dim \mathcal{N}/\mathcal{E}_0 \le 1$, $\mathcal{E}_0 = \mathcal{E}$ if $m = 3$ and $\mathcal{N} = \mathcal{E}$ if $m \ge 5$.

Theorem 1.2 (Goldberg-Green). *Suppose that V satisfies* Assumption 1.1 *and that H is of exceptional type. Then, if $m \geq 5$, $W_\pm$ are bounded in $L^p(\mathbb{R}^m)$ for $1 \leq p < m/2$. They are bounded in $L^p(\mathbb{R}^m)$ also for $1 \leq p < m$ if $\mathcal{N} = \mathcal{E}_0$ and for $1 \leq p < \infty$ if $\mathcal{N} = \mathcal{E}_1$.*

In this paper, we show following theorems which in particular prove the corresponding result for $m = 3$ and that the conditions $\mathcal{N} = \mathcal{E}_0$ and $\mathcal{N} = \mathcal{E}_1$ of Theorem 1.2 are also necessary for the stated properties of $W_\pm$ respectively. We write P, P_0 and P_1 for the orthogonal projections onto $\mathcal{E}$, $\mathcal{E}_0$ and $\mathcal{E}_1$ respectively. Because $(-\Delta)^{-1}V$ is a real operator, we may take the bases of $\mathcal{N}$, $\mathcal{E}_0$ and $\mathcal{E}_1$ which consist of real functions and P, P_0 and P_1 are real operators: For the conjugation $(\mathcal{C}u)(x) = \overline{u(x)}$,

$$\mathcal{C}^{-1}P\mathcal{C} = P, \quad \mathcal{C}^{-1}P_0\mathcal{C} = P, \quad \mathcal{C}^{-1}P_1\mathcal{C} = P_1. \tag{1.7}$$

We state results for $m = 3$, $m = 5$ and $m \geq 6$ separately.

Theorem 1.3. *Let $m = 3$. Suppose that V satisfies* Assumption 1.1 *and that H is of exceptional type. Then:*

(1) *$W_\pm$ are bounded in $L^p(\mathbb{R}^3)$ for $1 < p < 3$.*

(2) *For $3 < p < \infty$, there exists a constant C such that*

$$\|(W_\pm \pm a\varphi \otimes |D|^{-1}V\varphi + P)u\|_{L^p} \leq C\|u\|_{L^p}, \tag{1.8}$$

where φ is the real function defined by (3.13) *(the canonical resonance), $a = 4\pi i|\langle V, \varphi\rangle|^{-2}$ and P may be replaced by $P \ominus P_1$.*

(3) *If $W_\pm$ are bounded in $L^p(\mathbb{R}^3)$ for some $3 < p < \infty$, then $\mathcal{N} = \mathcal{E}_1$. In this case they are bounded in $L^p(\mathbb{R}^3)$ for all $1 < p < \infty$.*

Theorem 1.4. *Let $m = 5$. Suppose that V satisfies* Assumption 1.1 *and that H is of exceptional type. Then:*

(1) *$W_\pm$ are bounded in $L^p(\mathbb{R}^5)$ for $1 < p < 5/2$.*

(2) *For $5/2 < p < 5$, there exists a constant C such that*

$$\left\| \left(W_\pm \pm a_0(|D|^{-1}V\varphi) \otimes \varphi + \frac{P}{2} \right) u \right\|_{L^p} \leq C\|u\|_{L^p}, \tag{1.9}$$

where $\varphi = PV$, V being considered as a function, $a_0 = i/(24\pi^2)$ and P may be replaced by $P \ominus P_0$. If $W_\pm$ are bounded in $L^p(\mathbb{R}^5)$ for some $\frac{5}{2} < p < 5$, then $\mathcal{N} = \mathcal{E}_0$. In this case they are bounded in $L^p(\mathbb{R}^5)$ for all $1 < p < 5$.

(3) *By virtue of* (1) *and* (2)*, the condition $\mathcal{E} = \mathcal{E}_0$ is necessary for $W_\pm$ to be bounded in $L^p(\mathbb{R}^5)$ for some $p > 5$. Suppose $\mathcal{E} = \mathcal{E}_0$. Then,*

$$\|(W_\pm + P)\, u\|_{L^p} \leq C\|u\|_{L^p} \tag{1.10}$$

for a constant C, where $P = P_0$ may be replaced by $P_0 \ominus P_1$. If $W_\pm$ are bounded in $L^p(\mathbb{R}^5)$ for some $p > m$, then $\mathcal{N} = \mathcal{E}_1$. In this case they are bounded in $L^p(\mathbb{R}^5)$ for all $1 < p < \infty$.

Theorem 1.5. *Let $m \geq 6$. Suppose that V satisfies* Assumption 1.1 *and that H is of exceptional type. Then:*

(1) *$W_\pm$ are bounded in $L^p(\mathbb{R}^m)$ for $1 < p < m/2$.*

(2) *For $\frac{m}{2} < p < m$, there exists a constant $C > 0$ such that*

$$\|(W_\pm + D_m P)u\|_{L^p} \leq C_p \|u\|_{L^p}, \tag{1.11}$$

where P may be replaced by $P \ominus P_0$ and

$$D_m = \begin{cases} \dfrac{\Gamma\left(\frac{m-2}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{m-1}{2}\right)}, & m \text{ is odd}, \quad (1.12) \\ \dfrac{2^m\Gamma\left(\frac{m}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{m-1}{2}\right)} \displaystyle\int_1^\infty (x^2+1)^{-(m-1)} dx, & m \text{ is even}. \quad (1.13) \end{cases}$$

If $W_\pm$ are bounded in $L^p(\mathbb{R}^m)$ for some $m/2 < p < m$ then, $\mathcal{E} = \mathcal{E}_0$. In this case they are bounded in $L^p(\mathbb{R}^m)$ for all $1 < p < m$

(3) *Suppose $\mathcal{E} = \mathcal{E}_0$. Let $m < p < \infty$. Then, for a constant C_p,*

$$\|(W_\pm + P)u\| \leq C_p \|u\|_{L^p}, \tag{1.14}$$

where P may be replaced by $P_0 \ominus P_1$. If $W_\pm$ are bounded in $L^p(\mathbb{R}^m)$ for some $p > m$, then $\mathcal{E} = \mathcal{E}_1$. In this case they are bounded in $L^p(\mathbb{R}^m)$ for all $1 < p < \infty$.

Remark 1.6. (1) The integral in (1.13) may be computed explicitly:

$$\int_1^\infty (x^2+1)^{-(m-1)} dx = \frac{\Gamma\left(m-\frac{3}{2}\right)}{4\Gamma(m-1)} \left(\sqrt{\pi} - \sum_{j=1}^{m-2} \frac{\Gamma(j)2^{-j+1}}{\Gamma\left(j+\frac{1}{2}\right)}\right). \tag{1.15}$$

(2) There are examples of V such that $\mathcal{E}_1 = \mathcal{E}_0 \subsetneq \mathcal{N}$, $\mathcal{E}_1 \subsetneq \mathcal{E}_0 = \mathcal{N}$ and $\mathcal{E}_1 \subsetneq \mathcal{E}_0 \subsetneq \mathcal{N}$ (see Example 8.4 of [13]).
(3) Murata's result (Theorem 1.2 of [20]) also implies that, if $\mathcal{N} \neq 0$, $W_\pm$ are not in general bounded in $L^p(\mathbb{R}^m)$ for $p > 3$ if $m = 3$ and for $p > \frac{m}{2}$ if $m \geq 5$.

The rest of the paper is devoted to the proof of Theorems. In spite that substantial part of Theorems 1.4 and 1.5 overlaps with Theorem 1.2 and that they miss critically important L^1-boundedness, we present the proof of Theorems which is very different from the one by Goldberg and Green ([10]). Our proof heavily uses harmonic analysis machinery, which produces sharper results for larger p's, however, at the same time, prevents us from reaching end points

$p = 1$ and $p = \infty$. We prove the theorems only for W_- since conjugation changes the direction of time, viz. $\mathcal{C}^{-1}e^{-itH}\mathcal{C} = e^{itH}$, and

$$W_+ = \mathcal{C}^{-1}W_-\mathcal{C}. \tag{1.16}$$

We use the following notation and conventions: The ℓ-th derivative of $f(x)$, $x \in \mathbb{R}$ is denoted by $f^{(\ell)}(x)$. $\Sigma = \mathbb{S}^{m-1} = \{x\colon x_1^2 + \cdots + x_m^2 = 1\}$ is the unit sphere in $\mathbb{R}^m$ and $\omega_{m-1} = 2\pi^{\frac{m}{2}}/\Gamma\left(\frac{m}{2}\right)$ is its area. The coupling and the inner product are anti-linear with respect to the first component,

$$(u, v) = \langle u, v\rangle = \int_{\mathbb{R}^n} \overline{u(x)}v(x)dx,$$

in accordance with the interchangeable notaion for the rank 1 operator

$$|u\rangle\langle v| = u \otimes v \ : \ \phi \mapsto u\langle v, \phi\rangle.$$

This notation is used also when v is in a certain function space and u in its dual space.

$$f \leq_{|\,\cdot\,|} g \text{ means } |f| \leq |g|.$$

For Banach spaces X and Y, $\mathbf{B}(X, Y)$ is the Banach space of bounded operators from X to Y and $\mathbf{B}(X) = \mathbf{B}(X, X)$; $\mathbf{B}_\infty(X, Y)$ and $\mathbf{B}_\infty(X)$ are spaces of compact operators; and the dual space $\mathbf{B}(X, \mathbb{C})$ of X is denoted by X^*. The identity operators in various Banach spaces are indistinguishably denoted by 1. For $1 \leq p \leq \infty$, $\|u\|_p = \|u\|_{L^p}$ is the norm of $L^p(\mathbb{R}^m)$ and p' is its dual exponent, $1/p + 1/p' = 1$. When $p = 2$, we often omit p and write $\|u\|$ for $\|u\|_2$. We interchangeably write $L^p_w(\mathbb{R}^m)$ or $L^{p,\infty}(\mathbb{R}^m)$ for weak-L^p spaces and $\|u\|_{p,w}$ or $\|u\|_{p,\infty}$ for their norms. For $s \in \mathbb{R}$,

$$L^2_s = \langle x\rangle^{-s}L^2 = L^2(\mathbb{R}^m, \langle x\rangle^{2s}dx), \quad H^s(\mathbb{R}^m) = \mathcal{F}L^2_s(\mathbb{R}^m)$$

are the weighted L^2 spaces and Sobolev spaces. The space of rapidly decreasing functions is denoted by $\mathcal{S}(\mathbb{R}^m)$.

We denote the resolvents of H and H_0 respectively by

$$R(z) = (H - z)^{-1}, \quad R_0(z) = (H_0 - z)^{-1}.$$

We parameterize $z \in \mathbb{C} \setminus [0, \infty)$ as $z = \lambda^2$ by $\lambda \in \mathbb{C}^+$, the open upper half plane of $\mathbb{C}$, so that the positive and the negative parts of the boundary $\{\lambda\colon \pm\lambda \in (0, \infty)\}$ are mapped onto the upper and the lower edges of the positive half line $\{z \in \mathbb{C}\colon z > 0\}$. We define

$$G(\lambda) = R(\lambda^2), \quad G_0(\lambda) = R_0(\lambda^2), \quad \lambda \in \mathbb{C}^+.$$

These are $\mathbf{B}(\mathcal{H})$-valued meromorphic functions of $\lambda \in \mathbb{C}^+$ and the limiting absorption principle [19] (LAP for short) asserts that, when considered as $\mathbf{B}(\langle x\rangle^{-s}L^2, \langle x\rangle^{t}L^2)$-valued functions for $s, t > \frac{1}{2}$ and $s + t > 2$, $G_0(\lambda)$ has

Hölder continuous extensions to its closure $\overline{\mathbb{C}}^+ = \{z : \Im z \geq 0\}$. The same is true also for $G(\lambda)$, but, if H is of exceptional type, it has singularities at $\lambda = 0$. In what follows $z^{\frac{1}{2}}$ is the branch of square root of z cut along the negative real axis such that $z^{\frac{1}{2}} > 0$ when $z > 0$.

The plan of the paper is as follows: In section 2, we record some results most of which are well known and which we use in the sequel. They include:

- Formulas for the integral kernel of $G_0(\lambda)$ as exponential-polynomials in odd dimensions or their superpositions in even dimensions.
- Representation of $\langle\psi|(G_0(\lambda) - G_0(-\lambda))u\rangle$ as the linear combination of Fourier transforms of $r^{j+1}M(r, \overline{\psi} * \check{u})$, $M(r, f)$ being the average of f over the sphere of radius r centered at the origin.
- The Muckenhaupt weighted inequality and examples of A_p-weights.

In section 3, we recall and improve results of [33] and [8] on the behavior as $\lambda \to 0$ of $(1 + G_0(\lambda)V)^{-1}$ and reduce the problem to the L^p-boundedness of

$$Z_s u = -\frac{1}{\pi i}\int_0^\infty G_0(\lambda)VS(\lambda)(G_0(\lambda) - G_0(-\lambda))u\lambda F(\lambda)d\lambda \tag{1.17}$$

where $S(\lambda)$ is the singular part of the expansion of $(1 + G_0(\lambda)V)^{-1}$ at $\lambda = 0$ and $F \in C_0^\infty(\mathbb{R})$ is such that $F(\lambda) = 1$ near $\lambda = 0$.

We prove Theorem 1.3 in Section 4, Theorem 1.4 for odd dimensions $m \geq 5$ in Section 5 and for even dimensions in Section 6. We explain the basic strategy of the proof at the end of §4.1 after most of basic ideas appears in the simplest form.

Acknowledgment The part of this work was carried out while the author was visiting Aalborg and Aarhus universities in summers of 2014 and 2015 respectively. He would like to express his sincere gratitude to both institutions for the hospitality, to Arne Jensen, Jacob Schach Møller and Erik Skibsted in particular. He thanks Professors Goldberg and Green for sending us their manuscript [10] before posting on the arXiv-math, Shin Nakano for computing the recursion formula for (1.15) and Fumihiko Nakano for bringing [12] to his attention.

2 Preliminaries

In this section we record some well known results which we use in what follows.

2.1 Integral kernel of the free resolvent

For $m \geq 2$, resolvent $G_0(\lambda)$ for $\Im\lambda \geq 0$ is the convolution with

$$G_0(\lambda, x) = \frac{e^{i\lambda|x|}}{2(2\pi)^{\frac{m-1}{2}}\Gamma\left(\frac{m-1}{2}\right)|x|^{m-2}}\int_0^\infty e^{-t}t^{\frac{m-3}{2}}\left(\frac{t}{2} - i\lambda|x|\right)^{\frac{m-3}{2}} dt \tag{2.1}$$

([28]). When $m \geq 3$ is odd, it is an exponential polynomial like function.

Lemma 2.1. *Let $m \geq 3$ be odd. Then:*

$$G_0(\lambda, x) = \sum_{j=0}^{(m-3)/2} C_j \frac{(\lambda|x|)^j e^{i\lambda|x|}}{|x|^{m-2}} \ \textit{with}\ C_j = \frac{(-i)^j (m-3-j)!}{2^{m-1-j}\pi^{\frac{m-1}{2}} j! (\frac{m-3}{2}-j)!}. \tag{2.2}$$

The constant C_0 may also be written as $C_0 = (m-2)^{-1}\omega_{m-1}^{-1}$ and

$$iC_0 + C_1 = 0, \quad \textit{when } m \geq 5. \tag{2.3}$$

If m is even, the structure of $G_0(\lambda, x)$ is more complex and this makes the analysis harder. For partly circumventing the difficulty we express $G_0(\lambda, x)$ as a superposition of exponential-polynomial like functions of the form (2.2). This will allow a part of the proof for even dimensions to go in parallel with the odd dimensional cases. We set

$$\nu = \frac{m-2}{2}.$$

Define operators $T_j^{(a)}$, $j = 0, \ldots, \nu$ for superposing over parameter $a > 0$ by

$$T_j^{(a)}[f(x,a)] = C_{m,j}\omega_{m-1} \int_0^\infty (1+a)^{-(2\nu-j+\frac{1}{2})} f(x,a) \frac{da}{\sqrt{a}}, \tag{2.4}$$

$$C_{m,j}\omega_{m-1} = (-2i)^j \frac{\Gamma\left(2\nu - j + \frac{1}{2}\right)}{(m-2)!\sqrt{\pi}} \binom{\nu}{j}. \tag{2.5}$$

The factor ω_{m-1} is added for shorting some formulas below (see (2.18)).

Lemma 2.2. *If $m \geq 4$ is even, then we have*

$$G_0(\lambda, x) = \sum_{j=0}^{\nu} \omega_{m-1}^{-1} T_j^{(a)} \left[e^{i\lambda|x|(1+2a)} \frac{(\lambda|x|)^j}{|x|^{m-2}} \right]. \tag{2.6}$$

Proof. Let $C_{m*} = 2^{m-1}\pi^{\frac{m-1}{2}}\Gamma\left(\frac{m-1}{2}\right)$. In the formula (2.1):

$$G_0(\lambda, x) = \frac{e^{i\lambda|x|}}{C_{m*}|x|^{m-2}} \int_0^\infty e^{-t} t^{\frac{m-3}{2}} \left(t - 2i\lambda|x|\right)^{\frac{m-3}{2}} dt, \tag{2.7}$$

write $(t-2i\lambda|x|)^{\frac{m-3}{2}} = (t-2i\lambda|x|)^{\nu}(t-2i\lambda|x|)^{-\frac{1}{2}}$, expand $(t-2i\lambda|x|)^{\nu}$ via the binomial formula and use the identity

$$z^{-\frac{1}{2}} = \frac{1}{\sqrt{\pi}} \int_0^\infty e^{-az} a^{-\frac{1}{2}} da, \quad \Re z > 0 \tag{2.8}$$

for $(t - 2i\lambda|x|)^{-\frac{1}{2}}$. The right hand side of (2.7) becomes

$$\sum_{j=0}^{\nu} \frac{(-2i)^j}{\sqrt{\pi}C_{m*}} \binom{\nu}{j} \iint_{\mathbb{R}_+^2} e^{-(1+a)t} t^{2\nu-j} \left(e^{i\lambda|x|(1+2a)} \frac{(\lambda|x|)^j}{|x|^{m-2}} \right) \frac{dt}{\sqrt{t}} \frac{da}{\sqrt{a}}.$$

The integral converges absolutely if $m \geq 4$ and we obtain (2.6) after performing the integral with respect to t. $\square$

2.2 Spectral measure of H_0

. The spectral measure of $H_0 = -\Delta$ is AC and Stone's theorem implies that the spectral projection $E_0(d\mu)$ is given for $\mu = \lambda^2$, $\lambda > 0$ by

$$E_0(d\mu) = \frac{1}{2\pi i}(R_0(\mu + i0) - R_0(\mu - i0))d\mu = \frac{1}{i\pi}(G_0(\lambda) - G_0(-\lambda))\lambda d\lambda.$$

Lemma 2.3. *Let $m \geq 3$ and $u, v \in (L^1 \cap L^2)(\mathbb{R}^m)$. Then, both sides of the following equation can be continuously extended to $\lambda = 0$ and*

$$\lambda^{-1}\langle v, (G_0(\lambda) - G_0(-\lambda))u\rangle = \langle |D|^{-1}v, (G_0(\lambda) - G_0(-\lambda))u\rangle, \quad \lambda \geq 0. \tag{2.9}$$

For bounded continuous functions f on $\mathbb{R}$ we have for $\lambda \geq 0$,

$$f(\lambda)\langle v, (G_0(\lambda)u - G_0(-\lambda))u\rangle = \langle v, (G_0(\lambda)u - G_0(-\lambda))f(|D|)u\rangle. \tag{2.10}$$

Proof. For $u, v \in (L^1 \cap L^2)(\mathbb{R}^m)$ we have

$$\langle v, (G_0(\lambda) - G_0(-\lambda))u\rangle = \frac{\lambda^{m-2}i}{2(2\pi)^{m-1}}\int_\Sigma \overline{\hat{v}(\lambda\omega)}\hat{u}(\lambda\omega)d\omega, \tag{2.11}$$

where $\Sigma = \mathbb{S}^{m-1}$. It follows, since $\widehat{|D|^{-1}v}(\lambda\omega) = \lambda^{-1}\hat{v}(\lambda\omega)$, $\lambda > 0$, that

$$\langle |D|^{-1}v, (G_0(\lambda) - G_0(-\lambda))u\rangle = \frac{\lambda^{m-3}i}{2(2\pi)^{m-1}}\int_\Sigma \overline{\hat{v}(\lambda\omega)}\hat{u}(\lambda\omega)d\omega. \tag{2.12}$$

The right side extends to a continuous function of $\lambda \geq 0$ when $m \geq 3$ and (2.9) follows by comparing (2.11) and (2.12). Eqn. (2.10) likewise follows. $\square$

We define the spherical average of a function f on $\mathbb{R}^m$ by

$$M(r, f) = \frac{1}{\omega_{m-1}}\int_\Sigma f(r\omega)d\omega, \quad \text{for all } r \in \mathbb{R}. \tag{2.13}$$

We often write $M_f(r) = M(r, f)$. We have $M_f(-r) = M_f(r)$ and Hölder's inequality implies

$$\left(\frac{1}{\omega_{m-1}}\int_0^\infty |M_f(r)|^p r^{m-1}dr\right)^{1/p} \leq \|f\|_p, \quad 1 \leq p \leq \infty. \tag{2.14}$$

For an even function $M(r)$ of $r \in \mathbb{R}$, define $\tilde{M}(\rho)$ by

$$\tilde{M}(\rho) = \int_\rho^\infty rM(r)dr\left(= -\int_{-\infty}^\rho rM(r)dr\right). \tag{2.15}$$

Lemma 2.4. *Suppose $M(r) = M(-r)$ and $\langle r\rangle^2 M(r)$ is integrable. Then,*

$$\int_\mathbb{R} e^{-ir\lambda}rM(r)dr = \frac{\lambda}{i}\int_\mathbb{R} e^{-ir\lambda}\tilde{M}(r)dr, \ \int_\mathbb{R}\tilde{M}(r)dr = \int_\mathbb{R} r^2M(r)dr. \tag{2.16}$$

Proof. Since $rM(r) = -\tilde{M}(r)'$, integration by parts gives the first equation. We differentiate both sides of the first and set $\lambda = 0$. The second follows. $\square$

We denote $\breve{u}(x) = u(-x)$, $x \in \mathbb{R}^m$. (The sign $\breve{u}$ will be reserved for this purpose and will not be used to denote the conjugate Fourier transform.)

Representation formula for odd dimensions.

Lemma 2.5. *Let $m \geq 3$ be odd and $u, \psi \in C_0^\infty(\mathbb{R}^m)$. Define $c_j = \omega_{m-1} C_j$, $1 \leq j \leq \frac{m-3}{2}$, where C_j are the constants in* (2.2). *Then, for $\lambda > 0$ we have*

$$\langle \psi, (G_0(\lambda) - G_0(-\lambda))u \rangle = \sum_{j=0}^{\frac{m-3}{2}} c_j (-1)^{j+1} \lambda^j \int_{\mathbb{R}} e^{-i\lambda r} r^{1+j} M_{\overline{\psi} * \check{u}}(r) dr. \quad (2.17)$$

Proof. We compute $\langle \psi, G_0(\lambda)u \rangle$ by using the integral kernel (2.2) of $G_0(\lambda)$. Change the order of integration and use polar coordinates. Then,

$$\langle \psi, G_0(\lambda)u \rangle = \sum_{j=0}^{\frac{m-3}{2}} C_j \int_{\mathbb{R}^m} \overline{\psi(x)} \left(\int_{\mathbb{R}^m} \frac{\lambda^j e^{i\lambda|y|} u(x-y)}{|y|^{m-2-j}} dy \right) dx$$

$$= \sum_{j=0}^{\frac{m-3}{2}} C_j \int_{\mathbb{R}^m} \frac{\lambda^j e^{i\lambda|y|} (\overline{\psi} * \check{u})(y)}{|y|^{m-2-j}} dy = \sum_{j=0}^{\frac{m-3}{2}} c_j \int_0^\infty \lambda^j e^{i\lambda r} r^{1+j} M_{\overline{\psi} * \check{u}}(r) dr.$$

Since $M_{\overline{\psi} * \check{u}}(r)$ is even, change of variable r by $-r$ yields

$$-\langle \psi, G_0(-\lambda)u \rangle = \sum_{j=0}^{\frac{m-3}{2}} c_j \int_{-\infty}^0 \lambda^j e^{i\lambda r} r^{1+j} M_{\overline{\psi} * \check{u}}(r) dr.$$

Add both sides of last two equations and change r by $-r$. □

Representation formula for even dimensions. If m is even, we have the analogue of (2.17). For a function $M(r)$ on $\mathbb{R}$ and $a > 0$, define

$$M^a(r) = M((1+2a)^{-1} r).$$

Lemma 2.6. *Let $m \geq 2$. Let $u, \psi \in C_0^\infty(\mathbb{R}^m)$. Then*

$$\langle \psi, (G_0(\lambda) - G_0(-\lambda))u \rangle = \sum_{j=0}^{\nu} (-1)^{j+1} T_j^{(a)} \left[\frac{\lambda^j \mathcal{F}(r^{j+1} M^{a}_{\overline{\psi} * \check{u}})(\lambda)}{(1+2a)^{j+2}} \right], \quad (2.18)$$

$$\textit{For } j = 0, \ -T_0^{(a)} \left[\frac{\mathcal{F}(r M^{a}_{\overline{\psi} * \check{u}})(\lambda)}{(1+2a)^2} \right] = i T_0^{(a)} \left[\frac{\lambda (\mathcal{F} \widetilde{M^{a}_{\overline{\psi} * \check{u}}})(\lambda)}{(1+2a)^2} \right]. \quad (2.19)$$

Proof. Define $B_j(\lambda, r, a) = e^{i\lambda r(1+2a)} (\lambda r)^j r^{-(m-2)}$ and

$$B_j(\lambda, a)u(x) = \int_{\mathbb{R}^m} B_j(\lambda, |y|, a) u(x-y) dy, \quad j = 0, \dots, \nu.$$

Then, (2.6) and change of the order of integrations imply

$$\langle \psi, (G_0(\lambda) - G_0(-\lambda))u \rangle = \sum_{j=0}^{\nu} \frac{T_j^{(a)}}{\omega_{m-1}} \left[\langle \psi, (B_j(\lambda, a) - B_j(-\lambda, a))u \rangle \right]. \quad (2.20)$$

We have, as in odd dimensions, that for $u \in \mathcal{S}(\mathbb{R}^m)$ and $\psi \in L^1(\mathbb{R}^m)$

$$\langle \psi, B_j(\lambda,a)u\rangle = \int_{\mathbb{R}^m}\left(\int_{\mathbb{R}^m}\overline{\psi(x)}B_j(\lambda,|y|,a)u(x-y)dy\right)dx$$
$$= \int_{\mathbb{R}^m} B_j(\lambda,|y|,a)(\overline{\psi}*\check{u})(y)dy = \omega_{m-1}\int_0^\infty e^{i(1+2a)\lambda r}(\lambda r)^j r M_{\overline{\psi}*\check{u}}(r)dr.$$

Replacing λ to $-\lambda$ and changing the variable r to $-r$, we have

$$-\langle \psi, B_j(-\lambda,a)u\rangle = \omega_{m-1}\int_{-\infty}^0 e^{i(1+2a)\lambda r}(\lambda r)^j r M_{\overline{\psi}*\check{u}}(r)dr,$$

where we used that $M_{\overline{\psi}*\check{u}}(-r) = M_{\overline{\psi}*\check{u}}(r)$. Adding these two yields

$$\langle \psi, (B_j(\lambda,a)-B_j(-\lambda,a))u\rangle = \omega_{m-1}\int_{\mathbb{R}} e^{i(1+2a)\lambda r}(\lambda r)^j r M_{\overline{\psi}*\check{u}}(r)dr. \quad (2.21)$$

Change r to $-r$ in the right of (2.21), plug the result with (2.20) and, at the end, change the variable r to $-r/(1+2a)$. Then, (2.21) becomes

$$\frac{(-1)^{j+1}\omega_{m-1}}{(1+2a)^{j+2}}\int_{\mathbb{R}} e^{-i\lambda r}\lambda^j r^{j+1}M^{a}_{\overline{\psi}*\check{u}}(r)dr = \frac{(-1)^{j+1}\omega_{m-1}\lambda^j}{(1+2a)^{j+2}}\mathcal{F}(r^{j+1}M^{a}_{\overline{\psi}*\check{u}})(\lambda)$$

and (2.18) follows. If we use the first of (2.16), the right of the last equation for $j=0$ becomes

$$\frac{i\lambda(\mathcal{F}\widetilde{M^{a}_{\overline{\psi}*\check{u}}})(\lambda)}{(1+2a)^2}\omega_{m-1}$$

and we obtain (2.19). ☐

2.3 Some results from harmonic analysis.

The following lemma on weighted inequality (cf. [11], Chapter 9) plays crucial role in this paper. We let $1<p<\infty$ in this subsection.

Lemma 2.7. *The weight function $|r|^a$ is an A_p weight on $\mathbb{R}$ if and only if $-1<a<p-1$. The Hilbert transform $\tilde{\mathcal{H}}$ and the Hardy-Littlewood maximal operator $\mathcal{M}$ are bounded in $L^p(\mathbb{R}, w(r)dr)$ for A_p weights $w(r)$.*

Modifying the Hilbert transform $\tilde{\mathcal{H}}$, we define

$$\mathcal{H}u(\rho) = \frac{(1+\tilde{\mathcal{H}})u(\rho)}{2} = \frac{1}{2\pi}\int_0^\infty e^{ir\rho}\hat{u}(r)dr. \quad (2.22)$$

We shall repeatedly use following A_p weights on $\mathbb{R}^1$ to the operator $\mathcal{M}\mathcal{H}$:

$$|r|^{m-1-p(m-1)},\ |r|^{m-1-2p},\ |r|^{m-1-p} \ \text{ and } \ |r|^{m-1}, \quad (2.23)$$

respectively for $1 < p < \frac{m}{m-1}$, $\frac{m}{3} < p < \frac{m}{2}$, $\frac{m}{2} < p < m$ and $m < p$.
For a function $F(x)$ on $\mathbb{R}^m$, we say $G(|x|) \in L^1(\mathbb{R}^m)$ is a radially decreasing integrable majorant (RDIM for short) of F if $G(r) > 0$ is decreasing and $|F(x)| \leq G(|x|)$ for a.e. $x \in \mathbb{R}^m$. The following lemma is well known (see e.g. [26], p.57).

Lemma 2.8. (1) *A rapidly decreasing function $F \in \mathcal{S}(\mathbb{R}^m)$ has a RDIM.*

(2) *If F has a RDIM. then there is a constant $C > 0$ such that*

$$|(F * u)(t)| \leq C(\mathcal{M}u)(t), \quad t \in \mathbb{R}. \tag{2.24}$$

Lemma 2.9. *For u and $F \in L^1(\mathbb{R})$ such that $\hat{u}$, $\hat{F} \in L^1(\mathbb{R})$ we have*

$$\frac{1}{2\pi}\int_0^\infty e^{i\lambda\rho}F(\lambda)\hat{u}(\lambda)d\lambda = (\mathcal{F}^*F * \mathcal{H}u)(\rho). \tag{2.25}$$

Proof. Let $\Theta(\lambda) = \begin{cases} 1, & \text{for } \lambda > 0 \\ 0, & \text{for } \lambda \leq 0 \end{cases}$. Then, the left side of (2.25) equals

$$\begin{aligned}\frac{1}{2\pi}\int_{\mathbb{R}} e^{i\lambda\rho}F(\lambda)\Theta(\lambda)\hat{u}(\lambda)d\lambda &= \frac{1}{2\pi}\int_{\mathbb{R}}\left(\int_{\mathbb{R}} e^{i\lambda(\rho-\xi)}\mathcal{F}^*F(\xi)d\xi\right)\Theta(\lambda)\hat{u}(\lambda)d\lambda \\ &= \int_{\mathbb{R}} \mathcal{F}^*F(\xi)\mathcal{F}^*\{\Theta(\lambda)\hat{u}(\lambda)\}(\rho-\xi)d\xi = (\mathcal{F}^*F * \mathcal{H}u)(\rho)\end{aligned}$$

as desired. □

3 Reduction to the low energy analysis

We write $W_- = W$ in the sequel. When $u \in \langle x\rangle^{-s}L^2$, $s > 1/2$, Wu may be expressed via the boundary values of resolvents (e.g. [19]):

$$Wu = u - \lim_{\varepsilon\downarrow 0, N\uparrow\infty}\frac{1}{\pi i}\int_\varepsilon^N G(\lambda)V(G_0(\lambda) - G_0(-\lambda))u\lambda d\lambda \tag{3.1}$$

$$= u - \frac{1}{\pi i}\int_0^\infty G(\lambda)V(G_0(\lambda) - G_0(-\lambda))u\lambda d\lambda \tag{3.2}$$

Here the right of (3.1) is the Riemann integral of an $\langle x\rangle^t L^2$-valued continuous function where $t > 1/2$ is such that $s + t > 2$, the result belongs to $L^2(\mathbb{R}^m)$ and the limit exists in $L^2(\mathbb{R}^m)$, which we symbolically write as (3.2).
We decompose W into the high and the low energy parts

$$W = W_> + W_< \equiv W\Psi(H_0) + W\Phi(H_0), \tag{3.3}$$

by using cut off functions $\Phi \in C_0^\infty(\mathbb{R})$ and $\Psi \in C^\infty(\mathbb{R})$ such that

$$\Phi(\lambda^2) + \Psi(\lambda^2) \equiv 1, \quad \Phi(\lambda^2) = 1 \text{ near } \lambda = 0 \text{ and } \Phi(\lambda^2) = 0 \text{ for } |\lambda| > \lambda_0$$

for a small constant $\lambda_0 > 0$. We have proven in previous papers [33, 8] that, under Assumption 1.1, $W_>$ is bounded in $L^p(\mathbb{R}^m)$ for all $1 \le p \le \infty$ if $m \ge 3$ and we only need to study $W_< = \Phi(H_0) + Z$ where

$$Z = -\frac{1}{\pi i}\int_0^\infty G(\lambda)V(G_0(\lambda) - G_0(-\lambda))\lambda\Phi(H_0)d\lambda. \tag{3.4}$$

Evidently $\Phi(H_0) \in \mathbf{B}(L^p(\mathbb{R}^m))$ for all $1 \le p \le \infty$ and we have only to study the operator Z defined by (3.4). Since $\delta > 2$, the LAP (cf. Lemma 2.2 of [33]) implies that $G_0(\lambda)V$ is a Hölder continuous function of $\lambda \in \mathbb{R}$ with values in $\mathbf{B}_\infty(L^{-s})$ for any $\frac{1}{2} < s < \delta - \frac{1}{2}$ and, the absence of positive eigenvalues ([17]) implies that $1 + G_0(\lambda)V$ is invertible for $\lambda > 0$ (cf. [1]). It follows from the resolvent equation $G(\lambda) = G_0(\lambda) - G_0(\lambda)VG(\lambda)$ that $G(\lambda)V$ may be expressed in terms of $G_0(\lambda)V$:

$$G(\lambda)V = G_0(\lambda)V(1 + G_0(\lambda)V)^{-1} \text{ for } \lambda \ne 0 \tag{3.5}$$

and it is *locally* Hölder continuous for $\lambda \in \mathbb{R} \setminus \{0\}$ with values in $\mathbf{B}_\infty(L^2_{-s})$. Thus, we have the expression of Z in terms of the free resolvent $G_0(\lambda)$:

$$Zu = -\frac{1}{\pi i}\int_0^\infty G_0(\lambda)V(1 + G_0(\lambda)V)^{-1}(G_0(\lambda) - G_0(-\lambda))\lambda F(\lambda)ud\lambda, \tag{3.6}$$

where $F(\lambda) = \Phi(\lambda^2)$. If H is of generic type, $\mathrm{Ker}_{L^2_{-s}}(1 + G_0(0)V) = \mathcal{N} = \{0\}$ for any $\frac{1}{2} < s < \delta - \frac{1}{2}$ and $1 + G_0(\lambda)V$ is invertible for λ in a neighborhood $\lambda = 0$ and both sides of (3.5) become Hölder continuous. We then have shown in [33, 8] that Z is bounded in $L^p(\mathbb{R}^m)$ for all $1 \le p \le \infty$ under Assumption 1.1.

3.1 Low energy behavior of $(1 + G_0(\lambda)V)^{-1}$.

If H is of exceptional type, $(1 + G_0(\lambda)V)^{-1}$ becomes singular at $\lambda = 0$ and we describe its singularities here. Before doing so we recall some properties of functions in $\mathcal{N}$. Recall ([25]) that for $0 < s < m$:

$$|D|^{-s}u(x) = \mathcal{F}^*(|\xi|^{-s}\hat{u})(x) = \frac{\Gamma\left(\frac{m-s}{2}\right)}{2^s\pi^{\frac{m}{2}}\Gamma\left(\frac{s}{2}\right)}\int_{\mathbb{R}^m}\frac{u(y)}{|x-y|^{m-s}}dy. \tag{3.7}$$

When $s = 1$ and $s = 2$, the constants in front of the integral respectively equal to $\pi^{-1}\omega_{m-2}^{-1}$ and $C_0 = (m-2)^{-1}\omega_{m-1}^{-1}$ of (2.2).

Lemma 3.1. (1) *Functions ϕ in $\mathcal{N}$ satisfy $\langle x\rangle^{-s}\phi \in H^2(\mathbb{R}^m) \cap C^1(\mathbb{R}^m)$ for any $s > 1/2$ and $\nabla\phi$ is Hölder continuous. They satisfy the following asymptotic expansion as $|x| \to \infty$:*

$$\phi(x) = -\frac{C_0}{|x|^{m-2}}\int_{\mathbb{R}^m}(V\phi)(y)dy - \frac{1}{\omega_{m-1}}\sum_{j=1}^m \frac{x_j}{|x|^m}\int_{\mathbb{R}^m} y_j(V\phi)(y)dy + O(|x|^{-m}). \tag{3.8}$$

(2) *For* $\phi \in \mathcal{N} \setminus \mathcal{E}_0$, $\phi \otimes \phi \notin \mathbf{B}(L^p(\mathbb{R}^m))$ *for any* $1 \le p \le \infty$ *if* $m = 3$ *or* $m = 4$ *and, if* $m \ge 5$, $\phi \otimes \phi \in \mathbf{B}(L^p(\mathbb{R}^m))$ *if and only if* $\frac{m}{m-2} < p < \frac{m}{2}$. *If* $\phi \in \mathcal{E}_0 \setminus \mathcal{E}_1$, *then* $\phi \otimes \phi \in \mathbf{B}(L^p(\mathbb{R}^m))$ *if and only if* $\frac{m}{m-1} < p < m$ *for any* $m \ge 3$ *and, if* $\phi \in \mathcal{E}_1$, *then* $\phi \otimes \phi \in \mathbf{B}(L^p(\mathbb{R}^m))$ *for all* $1 < p < \infty$.

(3) *If* $\langle x\rangle^2 u \in L^1(\mathbb{R}^m)$, $|D|^{-1}u(x)$ *has the following expansion as* $|x| \to \infty$:

$$\frac{\int_{\mathbb{R}^m} u dx}{\pi\omega_{m-2}|x|^{m-1}} + \sum_{j=1}^{m} \frac{(m-1)x_j}{\pi\omega_{m-2}|x|^{m+1}} \int_{\mathbb{R}^m} x_j u dx + O(|x|^{-m-1}). \tag{3.9}$$

Proof. (1) The smoothness property of ϕ is well known (see e.g. Corollary 2.6 of [2]). We have from (3.7) that

$$\phi(x) = -C_0 \int_{\mathbb{R}^m} \frac{V(y)\phi(y)}{|x-y|^{m-2}} dy. \tag{3.10}$$

Taylor's formula implies that

$$\left| \frac{1}{|x-y|^{m-2}} - \frac{1}{|x|^{m-2}} - \frac{(m-2)x\cdot y}{|x|^{m-1}} \right| \le C \frac{\langle y\rangle^2}{\langle x\rangle^m}, \quad |x-y| \ge 1$$

and (3.8) follows. Statement (2) follows from (3.8). We omit the proof of (3) which is similar to that of (3.9). ☐

3.1.1 Odd dimensional cases

The structure of singularities depends on m. For odd dimensions $m \ge 3$ we have the following results (see, e.g. Theorem 2.12 of [33]). We state it separately for $m = 3$ and $m \ge 5$. In the following Theorems 3.2 and 3.3 for odd $m \ge 3$ and Theorem 3.4 for even $m \ge 6$, we will indiscriminately write $E(\lambda)$ for the operator valued function of λ defined near $\lambda = 0$ which, when inserted in (3.6) for $(1 + G_0(\lambda)V)^{-1}$, produces the operator which is bounded in $L^p(\mathbb{R}^m)$ for all $1 \le p \le \infty$.

The case $m = 3$. By virtue of (3.8), we have for $\phi \in \mathcal{N}$ that

$$\phi(x) = \frac{L(\phi)}{|x|} + O(|x|^{-2}) \text{ as } |x| \to \infty, \quad L(\phi) = \frac{-1}{4\pi} \int_{\mathbb{R}^3} V(x)\phi(x)dx. \tag{3.11}$$

Thus, $\mathcal{E} = \{\phi \in \mathcal{N} \setminus \{0\} \colon L(\phi) = 0\} (= \mathcal{E}_0)$ and, as $\mathcal{N} \ni \phi \mapsto L(\phi) \in \mathbb{C}$ is continuous, $\dim \mathcal{N}/\mathcal{E} \le 1$. Any $\varphi \in \mathcal{N} \setminus \mathcal{E}$ is called *threshold resonance* of H. We say that H is of exceptional type of *the first kind* if $\mathcal{E} = \{0\}$, *the second* if $\mathcal{E} = \mathcal{N}$ and *the third kind* if $\{0\} \subsetneq \mathcal{E} \subsetneq \mathcal{N}$. We let $D_0, D_1, \ldots$ be integral operators defined by

$$D_j u(x) = \frac{1}{4\pi j!} \int_{\mathbb{R}^3} |x-y|^{j-1} u(y) dy, \quad j = 0, 1, \ldots.$$

so that we have the formal Taylor expansion

$$G_0(\lambda)u(x) = \frac{1}{4\pi}\int_{\mathbb{R}^3} \frac{e^{i\lambda|x-y|}}{|x-y|}u(y)dy = \sum_{j=1}^{\infty}(i\lambda)^j D_j u.$$

If H is of exceptional type of the third kind, $-(V\phi,\phi)$ defines inner product on $\mathcal{N}$ and there is a unique real $\psi \in \mathcal{N}$ such that

$$-(V\psi,\phi)=0, \quad \forall\phi\in\mathcal{E}, \quad -(V\psi,\psi)=1 \text{ and } L(\psi)>0. \tag{3.12}$$

We define *the canonical resonance* by

$$\varphi = \psi + PVD_2V\psi \in \mathcal{N}. \tag{3.13}$$

If H is of exceptional type of the first kind, then $\dim\mathcal{N}=1$ and there is a unique $\varphi\in\mathcal{N}$ such that $-(V\varphi,\varphi)=1$ and $L(\varphi)>0$ and we call this the canonical resonance. We have the following result for $m=3$ (see e.g. [33]).

Theorem 3.2. *Let $m=3$ and let V satisfy $|V(x)|\le C\langle x\rangle^{-\delta}$ for some $\delta>3$. Suppose that H is of exceptional type of the third kind and let φ be the canonical resonance and $a=4\pi i|\langle V,\varphi\rangle|^{-2}$. Then:*

$$(I+G_0(\lambda)V)^{-1} = \frac{PV}{\lambda^2} + i\frac{PVD_3VPV}{\lambda} - \frac{a}{\lambda}|\varphi\rangle\langle\varphi|V + E(\lambda). \tag{3.14}$$

If H is of exceptional type of the first or the second kind, (3.14) *holds with $P=0$ or $\varphi=0$ respectively.*

The case $m\ge 5$. If $m\ge5$, (3.8) implies $\mathcal{N}=\mathcal{E}$.

Theorem 3.3. *Let $m\ge5$ be odd and $|V(x)|\le C\langle x\rangle^{-\delta}$ for some $\delta>m+3$. Suppose H is of exceptional type. Then:*

(1) *If $m=5$ then, with $\varphi=PV$, V being considered as a function,*

$$(I+G_0(\lambda)V)^{-1} = \frac{PV}{\lambda^2} - \frac{a_0}{\lambda}|\varphi\rangle\langle\varphi|V + E(\lambda), \quad a_0 = \frac{i}{24\pi^2}. \tag{3.15}$$

(2) *If $m\ge7$ then*

$$(I+G_0(\lambda)V)^{-1} = \frac{PV}{\lambda^2} + E(\lambda). \tag{3.16}$$

Define $S(\lambda)=(I+G_0(\lambda)V)^{-1}-E(\lambda)$ and

$$Z_s = \frac{i}{\pi}\int_0^\infty G_0(\lambda)VS(\lambda)(G_0(\lambda)-G_0(-\lambda))F(\lambda)\lambda d\lambda. \tag{3.17}$$

Then, it follows from Theorems 3.2 and 3.3 that $Z-Z_s\in\mathbf{B}(L^p(\mathbb{R}^m))$ for all $1\le p\le\infty$ and we have only to study Z_s in what follows.

3.2 Even dimensional case

When m is even, singular terms of $(1+G_0(\lambda)V)^{-1}$ may contain logarithmic factors. The following is the improvement of Proposition 3.6 of [8]. We let $\dim \mathcal{E} = d$ and $\{\phi_1, \ldots, \phi_d\}$ be the real orthonormal basis of $\mathcal{E}$. For making the expression simpler, we state the theorem for $V(1+G_0(\lambda)V)^{-1}$.

Theorem 3.4. *Let $m \geq 6$ be even. Suppose $|V(x)| \leq C\langle x\rangle^{-\delta}$ for $\delta > m+4$ if $m=6$ and for $\delta > m+3$ if $m \geq 8$. Let $\varphi = PV$ with V being considered as a function. Then, we have the following statements for $\Im\lambda \geq 0$ and $\log\lambda$ such that $\log\lambda \in \mathbb{R}$ for $\lambda > 0$:*

(1) *If $m = 6$ then, we have that*

$$V(1+G_0(\lambda)V)^{-1} = \frac{VPV}{\lambda^2} + \frac{\omega_5}{(2\pi)^6}\log\lambda(V\varphi\otimes V\varphi) + \left(\frac{\omega_5\|\varphi\|}{(2\pi)^6}\right)^2 \lambda^2\log^2\lambda(V\varphi\otimes V\varphi) + \lambda^2\log\lambda F_2 + VE(\lambda), \tag{3.18}$$

where F_2 is an operator of rank at most 8 such that

$$F_2 = \sum_{a,b=1}^{8} \varphi_a \otimes \psi_b, \quad \varphi_a,\ \psi_b \in (L^1\cap L^\infty)(\mathbb{R}^6). \tag{3.19}$$

(2) *If $m \geq 8$, then we have with a constant c_m that*

$$V(1+G_0(\lambda)V)^{-1} = \frac{VPV}{\lambda^2} + c_m(V\varphi\otimes V\varphi)\lambda^{m-6}\log\lambda + VE(\lambda). \tag{3.20}$$

(3) *If $m \geq 12$, then $c_m(V\varphi\otimes V\varphi)\lambda^{m-6}\log\lambda$ of (3.20) may be included in $VE(\lambda)$.*

Proof. We prove (1) only, using the notation of the proof of subsection 3.2.1 of [8]. A slightly more careful look at the argument there shows that, in spite of Eqn.(3.5) of [8], $V(1+G_0(\lambda)V)^{-1}$ is actually given by

$$\frac{VPV}{\lambda^2} + VD_{01}\log\lambda + VD_{21}\lambda^2\log\lambda + VD_{22}\lambda^2\log^2\lambda + VE(\lambda). \tag{3.21}$$

Here, with $F_{jk} = F_{jk}(0)$, $F_{jk}(\lambda)$ being defined by (3.16) of [8], and $A(0) = (2\pi)^{-6}\omega_{m-1}(1\otimes 1)$, VD_{01} and VD_{22} are rank 1 operators given by

$$VD_{01} = VPVF_{01}PV = VPVA(0)VPV = \frac{\omega_{m-1}}{(2\pi)^6}(V\varphi\otimes V\varphi), \tag{3.22}$$

$$VD_{22} = V(PVF_{01})^2PV = V(PVA(0)VP)^2V = \frac{\omega_{m-1}^2}{(2\pi)^{12}}\|\varphi\|^2(V\varphi\otimes V\varphi),$$

where we have used $PVQ = PV$ and $VQP = VP$ and,

$$VD_{21} = VPV(F_{21} + F_{00}PVF_{01} + F_{01}PVF_{00})PV \tag{3.23}$$
$$- VX(0)\overline{Q}D_2VPVF_{01}PV - VX(0)\overline{Q}A(0)VPV \tag{3.24}$$
$$- VPVF_{01}PVQD_2V\overline{Q}X(0) - PVQA(0)V\overline{Q}X(0). \tag{3.25}$$

It is obvious that the first line (3.23) is of rank at most 4 and of the form $\sum \alpha_{jk}(V\phi_j \otimes V\phi_k)$; four other operators are of rank one and of the form $f \otimes g$ with $f \in (L^1 \cap L^\infty)(\mathbb{R}^6)$. We check this for $VX(0)\overline{Q}D_2VPVF_{01}PV$ as a prototype. We have $D_2 = D_0^2$ and $D_0V\varphi = -\varphi$. Thus, (3.22) implies

$$VX(0)\overline{Q}D_2VPVF_{01}PV = -(2\pi)^{-6}\omega_{m-1}(VX(0)\overline{Q}D_0\varphi) \otimes (V\varphi).$$

Here $D_0\varphi \in C^2(\mathbb{R}^6)$ and satisfies $D_0\varphi \le_{|\,\cdot\,|} C\langle x\rangle^{-2}$ by virtue of Lemma 3.1. Hence, a fortiori $D_0\varphi \in C_0(\mathbb{R}^6)$, the Banach space of continuous functions which converge to 0 as $|x| \to \infty$. It is obvious that $\mathcal{X} \equiv \overline{Q}C_0(\mathbb{R}^6) \subset C_0(\mathbb{R}^6)$ and $X(0) = N^{-1}(0) = [\overline{Q}(1 + D_0V)\overline{Q}]^{-1}$ is an isomorphism of $\mathcal{X}$. This is because $T = \overline{Q}D_0V\overline{Q}$ is compact both in $\mathcal{X} = \overline{Q}C_0(\mathbb{R}^6)$ and $\mathcal{Y} = \overline{Q}L^2_{-\delta+2}(\mathbb{R}^6)$, $\mathcal{X} \cap \mathcal{Y}$ is dense in $\mathcal{Y}$ and $\mathrm{Ker}_{\mathcal{Y}}(1 + T) = \{0\}$ (see e.g. Lemma 2. 11 of [9]). Thus, $VX(0)\overline{Q}D_0\varphi(x) \le_{|\,\cdot\,|} C\langle x\rangle^{-\delta}$. □

It follows from Theorem 3.4 that $Zu = Z_su + Z_{\log}u$ modulo the operator which is bounded in L^p for all $1 \le p \le \infty$ and we need study

$$Z_{se} = \frac{i}{\pi}\int_0^\infty G_0(\lambda)VPV(G_0(\lambda) - G_0(-\lambda))F(\lambda)\lambda^{-1}d\lambda, \tag{3.26}$$

$$Z_{\log} = \sum_{j,k} \frac{i}{\pi}\int_0^\infty G_0(\lambda)\lambda^{2j}(\log\lambda)^k D_{jk}(G_0(\lambda) - G_0(-\lambda))F(\lambda)\lambda d\lambda, \tag{3.27}$$

for even $m \ge 6$, where the sum and D_{jk} are as in Theorem 3.4.

4 Proof of Theorem 1.3

The proof of Theorem 1.3 for $m = 3$ is the simplest and is the prototype for other dimensions and, most of the basic ideas already appear here.

4.1 The case of exceptional type of the first kind

We begin with the case that H is of exceptional type of the first kind and, we let φ be the canonical resonance, $a = 4\pi i|\langle V, \varphi\rangle|^{-2} \ne 0$ and

$$\psi(x) = |D|^{-1}(V\varphi)(x) = \frac{1}{2\pi^2}\int \frac{V(y)\varphi(y)}{|x-y|^2}dy. \tag{4.1}$$

The following lemma proves Theorem 1.3 when H is of exceptional type of the first kind.

Lemma 4.1. (1) *For $1 < p < 3$, there exists a constant C_p such that*

$$\|Z_s u\|_p \le C_p \|u\|_p, \quad u \in C_0^\infty(\mathbb{R}^3). \tag{4.2}$$

(2) *For $3 < p < \infty$, there exists a constant C_p such that*

$$\|(Z_s + a\varphi \otimes \psi)u\|_p \le C_p \|u\|_p, \quad u \in C_0^\infty(\mathbb{R}^3). \tag{4.3}$$

(3) *For $p \ge 3$, Z_s is unbounded in $L^p(\mathbb{R}^3)$.*

Proof. Recall $c_0 = C_0\omega_2 = 1$. We have $S(\lambda) = -\frac{a}{\lambda}|\varphi\rangle\langle\varphi|V$ and

$$Z_s u = -\frac{ia}{\pi}\int_0^\infty G_0(\lambda)V\varphi\rangle\langle V\varphi|(G_0(\lambda) - G_0(-\lambda))u\rangle F(\lambda)d\lambda. \tag{4.4}$$

Defining $M(r) = M(r, (V\varphi) * \check{u})$, we substitute (2.2) and (2.17) respectively for $G_0(\lambda)$ and $\langle V\varphi|(G_0(\lambda) - G_0(-\lambda))u\rangle$. Then,

$$Z_s u = \frac{ai}{\pi}\int_0^\infty \left(\int_{\mathbb{R}^3} \frac{e^{i\lambda|x-y|}V(y)\varphi(y)}{4\pi|x-y|}dy\right)\left(\int_{\mathbb{R}} e^{-i\lambda r} rM(r)dr\right)F(\lambda)d\lambda.$$

If we change the order of integrations,

$$Z_s u = \frac{ai}{2\pi}\int_{\mathbb{R}^3} \frac{K_0(|x-y|)V(y)\varphi(y)}{|x-y|}dy, \tag{4.5}$$

$$K_0(\rho) = \frac{1}{2\pi}\int_0^\infty e^{i\lambda\rho}F(\lambda)\left(\int_{\mathbb{R}} e^{-ir\lambda}rM(r)dr\right)d\lambda. \tag{4.6}$$

Since $\mathcal{F}^*F \in \mathcal{S}(\mathbb{R})$, it follows by virtue of Lemmas 2.8 and 2.9 that

$$K_0(\rho) = \{(\mathcal{F}^*F) * \mathcal{H}(rM(r))\}(\rho) \le_{|\cdot|} C\mathcal{MH}(rM)(\rho). \tag{4.7}$$

Function $K_0(\rho)$ may also be expressed as

$$K_0(\rho) = \frac{i}{2\pi\rho}\int_0^\infty e^{i\lambda\rho}\left(F(\lambda)\int_{\mathbb{R}} e^{-ir\lambda}rM(r)dr\right)' d\lambda. \tag{4.8}$$

and, after integration by parts, we see that $K_0(\rho)$ satisfies also

$$K_0(\rho) \le_{|\cdot|} C\rho^{-1}(\mathcal{MH}(r^2M)(\rho) + \mathcal{MH}(rM)(\rho)). \tag{4.9}$$

The boundary term does not appear in (4.8) since $\int_{\mathbb{R}} rM(r)dr = 0$.
(1a) Let $3/2 < p < 3$. By virtue of Young's inequality

$$\|Z_s u\|_p \le \frac{|a|(4\pi)^{1/p}}{2\pi}\|V\varphi\|_1 \left(\int_0^\infty \left|\frac{K_0(\rho)}{\rho}\right|^p \rho^2 d\rho\right)^{1/p}. \tag{4.10}$$

We estimate $K_0(\rho)$ by (4.7) and use that ρ^{2-p} is an A_p weight on $\mathbb{R}$. Lemma 2.7 and Young's inequality imply

$$\left(\int_0^\infty \left|\frac{K_0(\rho)}{\rho}\right|^p \rho^2 d\rho\right)^{1/p} \le C\left(\int_0^\infty |\mathcal{M}\mathcal{H}(rM)(\rho)|^p \rho^{2-p} d\rho\right)^{1/p}$$
$$\le C_p\left(\int_0^\infty M(r)^p r^2 dr\right)^{1/p} \le C_p\|V\varphi * u\|_p \le C_p\|V\varphi\|_1\|u\|_p. \tag{4.11}$$

and $\|Z_s u\|_p \le C_p\|V\varphi\|_1^2\|u\|_p$.
(1b) For $1 < p < \frac{3}{2}$, we use estimate (4.9) and that ρ^{2-2p} is an A_p weight on $\mathbb{R}$ and obtain that

$$\left(\int_0^\infty \left|\frac{K_0(\rho)}{\rho}\right|^p \rho^2 d\rho\right)^{\frac{1}{p}} \le \left(\int_0^\infty |(\mathcal{M}\mathcal{H}(r^2M) + \mathcal{M}\mathcal{H}(rM))(\rho)|^p \rho^{2-2p} d\rho\right)^{\frac{1}{p}}$$
$$\le C\left(\int_0^\infty |M(r)|^p \max(r^2, r^{2-p}) dr\right)^{\frac{1}{p}} \le C(\|V\varphi\|_1 + \|V\varphi\|_{p'})\|u\|_p, \tag{4.12}$$

where we estimated the integral over $0 \le r \le 1$ by using that

$$\sup|M(r)| \le \|V\varphi * u\|_\infty \le \|V\varphi\|_{p'}\|u\|_p. \tag{4.13}$$

Thus, we have $\|Z_s u\|_p \le C(\|V\varphi\|_1 + \|V\varphi\|_{p'})\|V\varphi\|_1\|u\|_p$ for $1 < p < 3/2$. Combining (1a) and (1b), we obtain (4.2) for $1 < p < 3$ by interpolation([5]).
(2) Let $p > 3$. Writing $\int_{\mathbb{R}} re^{-ir\lambda}M(r)dr = i\left(\int_{\mathbb{R}} e^{-ir\lambda}M(r)dr\right)'$ in (4.6), we apply integration by parts and obtain yet another expression of $K_0(\rho)$:

$$K_0(\rho) = \frac{-i}{2\pi}\int_{\mathbb{R}} M(r)dr - \frac{i}{2\pi}\int_0^\infty \left(e^{i\lambda\rho}F(\lambda)\right)'\left(\int_{\mathbb{R}} e^{-ir\lambda}M(r)dr\right)d\lambda. \tag{4.14}$$

Denote the second term by $\tilde{K}_0(\rho)$. By virtue of Lemmas 2.8 and 2.9,

$$\tilde{K}_0(\rho) \le_{|\cdot|} C(\rho+1)\mathcal{M}\mathcal{H}(M)(\rho). \tag{4.15}$$

Substituting (4.14) for $K_0(\rho)$ in (4.5), we obtain $Z_s u = Z_b u + Z_i u$, where Z_b and Z_i are operators produced by $\frac{-i}{2\pi}\int_{\mathbb{R}} M(r)dr$ and $\tilde{K}_0(\rho)$, respectively. Because

$$\frac{1}{\pi}\int_{\mathbb{R}} M(r)dr = \frac{1}{2\pi^2}\int_{\mathbb{R}^3}\left(\int_{\mathbb{R}^3}\frac{(V\varphi)(x+y)}{|x|^2}dx\right)u(y)dy = \langle\psi, u\rangle \tag{4.16}$$

by the definition (4.1), we have by using (3.10) for $m = 3$ that

$$Z_b u(x) = \frac{a}{4\pi^2}\int_{\mathbb{R}} M(r)dr \cdot \int_{\mathbb{R}^3}\frac{V(y)\varphi(y)}{|x-y|}dy = -a|\varphi\rangle\langle\psi|u\rangle. \tag{4.17}$$

We splite the integral as

$$Z_i u(x) = \frac{ai}{2\pi}\left(\int_{|y|\le 1} + \int_{|y|>1}\right)\frac{\tilde{K}_0(|y|)(V\varphi)(x-y)}{|y|}dy = I_1(x) + I_2(x). \tag{4.18}$$

For estimating I_2 we use (4.15) for $\rho \geq 1$: $|\tilde{K}_0(\rho)| \leq C\rho\mathcal{M}\mathcal{H}(M)(\rho)$. Since ρ^2 is an A_p-weight on $\mathbb{R}$ for $p > 3$, we have by using Young's and Hölder's inequalities and Lemma 2.7 that

$$\|I_2\|_p \leq C\|V\varphi\|_1 \left(\int_0^\infty |\mathcal{M}\mathcal{H}(M)(\rho)|^p \rho^2 d\rho\right)^{\frac{1}{p}}$$
$$\leq C\|V\varphi\|_1 \left(\int_0^\infty |M(r)|^p r^2 dr\right)^{\frac{1}{p}} \leq C\|V\varphi\|_1^2 \|u\|_p. \tag{4.19}$$

Hölder's inequality implies, with $p' = p/p-1$, that

$$|I_1(x)| \leq C \left(\int_{|y|\leq 1} \left|\frac{(V\varphi)(x-y)}{|y|}\right|^{p'} dy\right)^{1/p'} \left(\int_0^1 |\tilde{K}_0(\rho)|^p \rho^2 d\rho\right)^{1/p}.$$

Since $\tilde{K}_0(\rho) \leq_{|\,\cdot\,|} C\mathcal{M}\mathcal{H}(M)(\rho)$ for $0 < \rho < 1$ by virtue of (4.15) and since ρ^2 is an A_p-weight, we obtain as in (4.19) that

$$\left(\int_0^1 |\tilde{K}_0(\rho)|^p \rho^2 d\rho\right)^{1/p} \leq C \left(\int_0^\infty |\mathcal{M}\mathcal{H}(M)(\rho)|^p \rho^2 d\rho\right)^{1/p} \leq C\|u\|_p. \tag{4.20}$$

It follows by virtue of Minkowski's inequality that

$$\|I_1\|_p \leq C\|u\|_p \left\| \left(\int_{|y|\leq 1} \left|\frac{(V\varphi)(x-y)}{|y|}\right|^{p'} dy\right)^{1/p'} \right\|_p \leq C\|u\|_p\|V\varphi\|_p \tag{4.21}$$

because $1 < p' < 3/2 < 3 < p < \infty$. Thus,

$$\left\| \int_{\mathbb{R}^3} \frac{\tilde{K}_0(|x-y|)V(y)\varphi(y)}{|x-y|} dy \right\|_p \leq C(\|V\varphi\|_p + \|V\varphi\|_1)\|u\|_p.$$

With (4.17) this proves (4.3).
(3) Since $\int_{\mathbb{R}^3} V\varphi dx \neq 0$, Lemma 3.1 implies that $\varphi \notin L^p(\mathbb{R}^3)$ for $1 \leq p \leq 3$ and that $\psi \in L^p(\mathbb{R}^3)^*$ if and only if $p > 3$. Hence, $\varphi \otimes \psi$ is unbounded in $L^p(\mathbb{R}^3)$ for any $1 \leq p \leq \infty$. Thus, statement (2) implies that Z_s is unbounded in $L^p(\mathbb{R}^3)$ for $p \geq 3$. This completes the proof of the lemma. □

We review here the basic strategy of this subsection as it will be repeatedly employed in the following (sub)sections. We express $Z_s u$ as the convolution (4.5) of $V\varphi$ and $K_0(\rho)$ of (4.6). By applying integration by parts if necessary we represent and estimate $K_0(\rho)$ as in (4.7), (4.9) or (4.15) by using $\mathcal{M}\mathcal{H}$. These estimates are used for proving

$$\left(\int_0^\infty |K_0(\rho)|^p \rho^{2-p} d\rho\right)^{\frac{1}{p}} \left(= \omega_2^{-\frac{1}{p}} \left\|\frac{K_0(|x|)}{|x|}\right\|_p\right) \leq C\|u\|_p \tag{4.22}$$

via the weighted inequality for $\frac{3}{2} < p < 3$, $1 < p < \frac{3}{2}$ and $p > 3$ respectively. Desired estimates are then obtained by combining (4.22) and Young's inequality. However, the boundary term appears in the integration by parts for large values of $p > 3$ which obstructs the L^p-boundedness. We represent the obstruction explicitly in terms of functions of $\mathcal{N}$ and show that L^p-boundedness depends on the properties of functions in $\mathcal{N}$. Suitable modifications, improvements and additional arguments will be of course necessary in what follows.

4.2 The cases of the second and third kinds

Let H be of exceptional type of the second kind. Then,

$$S(\lambda) = \frac{PV}{\lambda^2} + i\frac{PVD_3VPV}{\lambda}, \tag{4.23}$$

where D_3 is the integral operator with kernel $|x-y|^2/4\pi$. We take the real orthonormal basis $\{\phi_1, \ldots, \phi_n\}$ of $\mathcal{E}$ and define $a_{jk} = \pi^{-1}\langle\phi_j|VD_3V|\phi_k\rangle \in \mathbb{R}$. We have $\langle V, \phi_j\rangle = 0$, $1 \le j \le n$. Substituting (4.23) for $S(\lambda)$ in (3.17), we have

$$Z_s u = Z_{s0}u + Z_{s1}u = \sum_{j,k=1}^{n} Z_{s0,jk}u + \sum_{j=1}^{n} Z_{s1,j}, \tag{4.24}$$

$$Z_{s0,jk}u = ia_{jk}\int_0^\infty G_0(\lambda)V\phi_j\rangle\langle V\phi_k|(G_0(\lambda)-G_0(-\lambda))u\rangle F(\lambda)d\lambda, \tag{4.25}$$

$$Z_{s1,j}u = \frac{i}{\pi}\int_0^\infty G_0(\lambda)V\phi_j\rangle\langle V\phi_j|(G_0(\lambda)-G_0(-\lambda))u\rangle F(\lambda)\frac{d\lambda}{\lambda}. \tag{4.26}$$

Lemma 4.2. *For any $1 < p < \infty$, there exists a constant C_p such that*

$$\|Z_{s0}u\|_p \le C_p\|u\|_p, \quad u \in C_0^\infty(\mathbb{R}^3). \tag{4.27}$$

Proof. The operator $Z_{s0,jk}$ is equal to Z_s of (4.4) with two $\varphi \in \mathcal{N}$'s being replaced by ϕ_j and $\phi_k \in \mathcal{E}$ and a by $-\pi a_{jk}$. Thus, the proof of Lemma 4.1 implies that $Z_{s0,jk} \in \mathbf{B}(L^p(\mathbb{R}^3))$ for $1 < p < 3$ and that

$$Z_{s0,jk} - \pi a_{jk}\phi_j \otimes |D|^{-1}(V\phi_k) \in \mathbf{B}(L^p(\mathbb{R}^3)), \quad p > 3. \tag{4.28}$$

Here $\phi_j \otimes |D|^{-1}(V\phi_k)$ is bounded in $L^p(\mathbb{R}^3)$ for $p > 3$ because $\phi_j \in L^p(\mathbb{R}^3)$ and $|D|^{-1}(V\phi_k) \in (L^p(\mathbb{R}^3))^*$ by virtue of (3.8) and (3.9). Thus $Z_{s0,jk} \in \mathbf{B}(L^p(\mathbb{R}^3))$ for $3 < p$ and, hence, for $1 < p < \infty$ by interpolation. This proves the lemma. ☐

Lemma 4.3. (1) *Let $1 < p < 3$. Then, for a constant C_p, we have*

$$\|Z_{s1}u\|_p \le C_p\|u\|_p, \quad u \in C_0^\infty(\mathbb{R}^3). \tag{4.29}$$

(2) *Let $3 < p < \infty$. Then, for a constant C_p, we have*

$$\|(Z_{s1}+P)u\|_p \le C\|u\|_p, \quad u \in C_0^\infty(\mathbb{R}^3). \tag{4.30}$$

In (4.30) *P may be replaced by $P \ominus P_1$ by virtue of* Lemma 3.1.

(3) *The operator Z_{s1} is bounded in $L^p(\mathbb{R}^3)$ for some $p > 3$ if and only if $\mathcal{E} = \mathcal{E}_1$. In this case Z_{s1} is bounded in $L^p(\mathbb{R}^3)$ for all $1 < p < \infty$.*

Proof. Define $\psi_j(x) = |D|^{-1}(V\phi_j)(x)$, $j = 1, \dots, n$. Then Lemma 2.3 implies

$$Z_{s1,j}u = \frac{i}{\pi}\int_0^\infty G_0(\lambda)|V\phi_j\rangle\langle\psi_j|(G_0(\lambda) - G_0(-\lambda))u\rangle F(\lambda)d\lambda \tag{4.31}$$

which can be obtained from $Z_s u$ of (4.4) by replacing a by -1, the first $V\varphi$ by $V\phi_j$ and the second by ψ_j. Thus, it may be expressed by using $K_{0,j}(\rho)$ of (4.6) with $M(r)$ being replaced by $M_j(r) = M(r, \psi_j * \check{u})$:

$$Z_{s1,j}u = \frac{1}{2\pi i}\int_{\mathbb{R}^3}\frac{K_{0,j}(|x-y|)V(y)\phi_j(y)}{|x-y|}dy. \tag{4.32}$$

(1) The argument of (1a) in the proof of Lemma 4.1 implies

$$\|Z_{s1,j}u\|_p \le C\|V\phi_j\|_1\|\psi_j * u\|_p, \quad 3/2 < p < 3 \tag{4.33}$$

(see (4.11)) and the one of (1b) does

$$\|Z_{s1,j}u\|_p \le C\|V\phi_j\|_1(\|\psi_j * u\|_p + \|\psi_j * u\|_\infty), \quad 1 < p < 3/2 \tag{4.34}$$

(see (4.12)). Since $\int V\phi_j dx = 0$, (3.9) implies that $\psi_j = |D|^{-1}\phi_j \in L^q(\mathbb{R}^3)$ for all $1 < q \le \infty$ and that the convolution operator with $\psi_j(x)$ is bounded in L^p for any $1 < p < \infty$ via Calderón-Zygmund theory (see e.g. [26], pp. 30-36). Thus, $\|\psi_j * u\|_p \le C\|u\|_p$, $\|\psi_j * u\|_\infty \le \|\psi_j\|_{p'}\|u\|_p$ and $Z_{s1,j}$ is bounded in $L^p(\mathbb{R}^3)$ for all $1 < p < 3$, $j = 1, \dots, n$. Statement (1) follows.
(2) Integration by parts as in (4.14) by using the identity $\int_{\mathbb{R}} e^{-ir\lambda} r M_j(r)dr = i\left(\int_{\mathbb{R}} e^{-ir\lambda}M_j(r)dr\right)'$ implies that $K_{0,j}(\rho)$ may be written as

$$-\frac{i}{2\pi}\int_{\mathbb{R}} M_j(r)dr - \frac{i}{2\pi}\int_0^\infty \left(e^{i\lambda\rho}F(\lambda)\right)'\left(\int_{\mathbb{R}} e^{-ir\lambda}M_j(r)dr\right)d\lambda, \tag{4.35}$$

which we insert into (4.32). Since $|D|^{-1}\psi_j = (-\Delta)^{-1}(V\phi_j) = -\phi_j$, (4.16) with $\psi_j \in \mathcal{E}$ in place of $V\varphi$ produces $-\langle\phi_j|u\rangle$. It follows that the boundary term of (4.35) produces

$$\frac{-1}{4\pi}\int_{\mathbb{R}^3}\frac{V(y)\phi_j(y)}{|x-y|}dy \cdot \frac{1}{\pi}\int_{\mathbb{R}} M_j(r)dr = -|\phi_j\rangle\langle\phi_j|u\rangle \tag{4.36}$$

as in (4.17). Denote by $\tilde{K}_{0j}(\rho)$ and $\tilde{Z}_{s1,j}$ the second term of (4.35) and the operator it produces via (4.32). They can respectively be obtained from $\tilde{K}_0(\rho)$ of (4.14) and Z_i of (4.18) by replacing $M(r)$ and $\tilde{K}_0(\rho)$ by $M_j(r)$ and $\tilde{K}_{0,j}(\rho)$. Thus, the argument of step (2) of the proof of Lemma 4.1, (4.19) and (4.21) in particular, implies that

$$\|\tilde{Z}_{s1,j}u\|_p \le C(\|V\phi_j\|_p + \|V\phi_j\|_1)\|\psi_j * u\|_p, \quad 3 < p < \infty. \tag{4.37}$$

The Calderón-Zygmund theory with (3.9) once more implies $\left\|\tilde{Z}_{s1,j}u\right\|_p \leq C\|u\|_p$. Since $\phi \otimes \phi \in \mathbf{B}(L^p)$ for all $1 < p < \infty$ if $\phi \in \mathcal{E}_1$ by virtue of (3.8), this together with (4.36) proves statement (2).
(3) It is obvious from (1) and (2) that $Z_{s1} \in \mathbf{B}(L^p(\mathbb{R}^3))$ for all $1 < p < \infty$ if $\mathcal{E} = \mathcal{E}_1$. Suppose then that $Z_{s1} \in \mathbf{B}(L^p(\mathbb{R}^3))$ for some $p > 3$ then $P \ominus P_1$ must be bounded in $L^p(\mathbb{R}^3)$ by virtue of (2). Take the orthonormal basis $\{\phi_1, \dots, \phi_d\}$ of $\mathcal{E} \ominus \mathcal{E}_1$ and $\{\rho_1, \dots, \rho_d\} \subset C_0^\infty(\mathbb{R}^3)$ such that $\{(\rho_j, \phi_k)\}$ becomes the unit matrix. Then, $(P \ominus P_1)\rho_j = \phi_j$, $j = 1, \dots, n$ and, if $P \ominus P_1$ is bounded in $L^p(\mathbb{R}^3)$ for some $p \geq 3$, there must exist a constant $C > 0$ such that

$$|(u, \phi_j)| = |((P \ominus P_1)u, \rho_j)| \leq C_j \|u\|_p, \quad \text{for all } u \in C_0^\infty(\mathbb{R}^3).$$

Then, ϕ_j has to be in $L^{p'}(\mathbb{R}^3)$ for $p' \leq 3/2$ for all $j = 1, \dots, n$. This implies $\phi_j = 0$ by virtue of (3.8). Thus, $\mathcal{E} = \mathcal{E}_1$ must hold. This completes the proof. □

Lemma 4.2 and Lemma 4.3 prove Theorem 1.3 when H is of exceptional type of the second kind. The following lemma completes the proof of Theorem 1.3.

Lemma 4.4. *Suppose that H is of exceptional type of the third kind. Then:*

(1) *W is bounded in $L^p(\mathbb{R}^3)$ for all $1 < p < 3$.*

(2) *$W + a\varphi \otimes (|D|^{-1}V\varphi) + P$ is bounded in $L^p(\mathbb{R}^3)$ for all $p > 3$.*

(3) *W is unbounded in $L^p(\mathbb{R}^3)$ for any $p > 3$ and $p = 1$.*

Proof. The combination of Lemmas 4.1, 4.2 and 4.3 proves statements (1) and (2). Suppose that W is bounded in $L^p(\mathbb{R}^3)$ for some $3 < p < \infty$. Then, so is $a(\varphi \otimes (|D|^{-1}V\varphi)) + P$. Let $\psi \in \mathcal{N}$ be the function which defines the canonical resonance φ by (3.13) and which satisfies (3.12). Then,

$$(V\psi, a(\varphi \otimes (|D|^{-1}V\varphi))u + Pu) = -a(|D|^{-1}V\varphi, u), \quad u \in C_0^\infty(\mathbb{R}^3)$$

and this must be extended to a bounded functional of $u \in L^p(\mathbb{R}^3)$. Hence, $|D|^{-1}V\varphi \in L^q(\mathbb{R}^3)$ for $q = (p-1)/p < 3/2$. This contradicts (3.8) because $\int_{\mathbb{R}^3} V(x)\varphi(x)dx \neq 0$ and (3) is proved. □

5 Proof of Theorems 1.4 and 1.5 for odd m

If $m \geq 5$, then $\mathcal{N} = \mathcal{E}$ and we let $\{\phi_1, \dots, \phi_d\}$ be the real orthonormal basis of $\mathcal{E}$. Theorem 3.3 implies that, with $a_0 = i/(24\pi^2)$,

$$S(\lambda) = \begin{cases} \lambda^{-2}PV - a_0\lambda^{-1}(\varphi \otimes V\varphi), & \text{if } m = 5, \\ \lambda^{-2}PV, & \text{if } m \geq 7. \end{cases} \tag{5.1}$$

Note that $\varphi \neq 0$ if and only if $\mathcal{E}_1 \neq \mathcal{E}$. We substitute (5.1) for $S(\lambda)$ in (3.17) and apply (2.2) and (2.17) as previously. Let C_j, c_k, $1 \le j,k \le \frac{m-3}{2}$ respectively be constants of (2.2) and (2.17). Then, we have

$$Z_s u = Z_{s0}u + Z_{s1}u, \tag{5.2}$$

where $Z_{s0} = 0$ for $m \ge 7$ and, for $m = 5$, with $M(r) = M(r, V\varphi * \check{u})$

$$Z_{s0}u = -2ia_0 \sum_{j,k=0,1} (-1)^{j+1} C_k c_j Z_{s0}^{jk} u, \tag{5.3}$$

$$Z_{s0}^{jk}u(x) = \int_{\mathbb{R}^5} \frac{V\varphi(y)}{|x-y|^{3-k}} K_0^{(j,k)}(|x-y|)dy, \tag{5.4}$$

$$K_0^{(j,k)}(\rho) = \frac{1}{2\pi}\int_0^\infty e^{i\lambda\rho}\lambda^{j+k}\left(\int_{\mathbb{R}} e^{-i\lambda r} r^{j+1} M(r)dr\right) F(\lambda)d\lambda, \tag{5.5}$$

and $Z_{s1}u$ is defined for all $m \ge 5$ by

$$Z_{s1}u = \sum_{l=1}^{d} Z_{s1}(\phi_l)u \tag{5.6}$$

where, for $\phi \in \mathcal{E}$, with $M(r) = M(r, V\phi * \check{u})$,

$$Z_{s1}(\phi)u = 2i \sum_{j,k=0}^{\frac{m-3}{2}} (-1)^{j+1} C_k c_j Z_{s1}^{jk}(\phi), \tag{5.7}$$

$$Z_{s1}^{jk}(\phi)u(x) = \int_{\mathbb{R}^m} \frac{V\phi(y)}{|x-y|^{m-2-k}} K^{(j,k)}(|x-y|)dy, \tag{5.8}$$

$$K^{(j,k)}(\rho) = \frac{1}{2\pi}\int_0^\infty e^{i\lambda\rho}\lambda^{j+k-1}\left(\int_{\mathbb{R}} e^{-i\lambda r} r^{j+1} M(r)dr\right) F(\lambda)d\lambda. \tag{5.9}$$

Note that $Z_{s0}^{jk}u$ and $K_0^{(j,k)}(\rho)$ are obtained from $Z_{s1}^{jk}u$ and $K^{(j,k)}(\rho)$ by changing ϕ by φ and λ^{j+k-1} by λ^{j+k} in (5.9).
We shall prove the last statements of (2) and (3) of Theorems 1.4 and 1.5 only for $Z_{s1}(\phi)$ since the proof of Lemma 4.3 (3) can easily be adapted for proving the same statements for Z_{s0}.

5.1 Estimate of Z_{s0} for $m = 5$

We begin by proving the following lemma for Z_{s0}, assuming $\varphi \neq 0$.

Lemma 5.1. (1) *Z_{s0} is bounded in $L^p(\mathbb{R}^5)$ for $1 < p < 5$.*

(2) *$Z_{s0} + a_0|\varphi\rangle\langle|D|^{-1}(V\varphi)|$ is bounded in $L^p(\mathbb{R}^5)$ for $5/2 < p < \infty$.*

(3) *Z_{s0} is not bounded in $L^p(\mathbb{R}^5)$ if $p \ge 5$.*

Proof. For $\varphi = PV$, we have $\int_{\mathbb{R}^5} V\varphi dx = \|\varphi\|^2 > 0$ and, by virtue of (3.8) and (3.9), $\varphi \otimes |D|^{-1}(V\varphi) \in \mathbf{B}(L^p(\mathbb{R}^5))$ if and only if $5/3 < p < 5$. Hence, statement (3) follows (2). Using that $e^{i\rho\lambda} = (i\rho)^{-(k+1)}\partial_\lambda^{k+1} e^{i\rho\lambda}$ and $\int_{\mathbb{R}} \lambda^{j+1} e^{-i\lambda r} M(r)dr = i^j \left(\int_{\mathbb{R}} e^{-i\lambda r} M(r)dr\right)^{(j)}$, we apply integration by parts to (5.5) and write $K_0^{(j,k)}(\rho)$ in two ways

$$K_0^{(j,k)}(\rho) = \frac{i^{k+1}}{2\pi\rho^{k+1}} \int_0^\infty e^{i\rho\lambda} \left(\lambda^{j+k} F(\lambda) \int_{\mathbb{R}} e^{-i\lambda r} r^{j+1} M(r)dr\right)^{(k+1)} d\lambda \tag{5.10}$$

$$= \frac{(-i)^j}{2\pi} \int_0^\infty \left(e^{i\lambda\rho} \lambda^{j+k} F(\lambda)\right)^{(j)} \left(\int_{\mathbb{R}} e^{-i\lambda r} rM(r)dr\right) d\lambda. \tag{5.11}$$

Note that boundary terms do not appear in (5.10) since $\int_{\mathbb{R}} rM(r)dr = 0$ and, if $k = 1$, we may apply further integration by parts to (5.11) without having boundary term and

$$K_0^{(j,k)}(\rho) = \frac{(-i)^{j+1}}{2\pi} \int_0^\infty \left(e^{i\lambda\rho} \lambda^{j+k} F(\lambda)\right)^{(j+1)} \left(\int_{\mathbb{R}} e^{-i\lambda r} M(r)dr\right) d\lambda. \tag{5.12}$$

We then apply Lemmas 2.8 and 2.9 to the right sides and obtain the following estimates for $j, k = 0, 1$:

$$K_0^{(j,k)}(\rho) \le_{|\cdot|} \begin{cases} C\rho^{-(k+1)} \sum_{l=0}^{k+1} \mathcal{M}\mathcal{H}(r^{j+l+1}M)(\rho), & (5.13) \\ C(1+\rho^{j+k})\mathcal{M}\mathcal{H}(r^{1-k}M)(\rho). & (5.14) \end{cases}$$

(a) Let $1 < p < 5/4$. Since $|r|^{-4(p-1)}$ is an A_p weight on $\mathbb{R}$ and $3p - 4 > -1$, we have by using (5.13) and (4.13) that, for any $j, k = 0, 1$,

$$\begin{aligned} \left\| \frac{K_0^{(j,k)}(|y|)}{|y|^{3-k}} \right\|_p &\le C \sum_{l=0}^{k+1} \left(\int_0^\infty \frac{|\mathcal{M}\mathcal{H}(r^{j+l+1}M)(\rho)|^p}{\rho^{4(p-1)}} d\rho\right)^{1/p} \\ &\le C \left(\int_0^1 \frac{|M(r)|^p dr}{r^{3p-4}} + \int_1^\infty |M(r)|^p r^4 dr\right)^{\frac{1}{p}} \le C(\|V\varphi\|_{p'} + \|V\varphi\|_1)\|u\|_p. \end{aligned} \tag{5.15}$$

Young's inequality then implies $\|Z_{s0}^{jk} u\|_p \le C\|V\varphi\|_1(\|V\varphi\|_{p'} + \|V\varphi\|_1)\|u\|_p$.
(b) We next show that $\|Z_{s0}^{j1} u\|_p \le C\|u\|_p$ for $p > 5$ and $j = 0, 1$. Interpolating this with the result of (a), we then have the same for all $1 < p < \infty$. We split the integral as in (4.18) and repeat the argument after it:

$$|Z_{0s}^{j1} u(x)| \le C \left(\int_{|y|\le 1} + \int_{|y|>1}\right) \frac{|V\varphi(x-y)|}{|y|^2} |K_0^{(j,1)}(|y|)| dy = I_1(x) + I_2(x).$$

For $\rho \geq 1$, we have $K_0^{(j,1)}(\rho) \leq_{|\cdot|} C\rho^2 \mathcal{M}\mathcal{H}(M(r))(\rho)$ by virtue of (5.14) and since r^4 is A_p weight on $\mathbb{R}$ if $p > 5$. It follows that

$$\|I_2\|_p \leq C\|V\varphi\|_1 \left\| \frac{K_0^{(j,1)}}{|x|^2} \right\|_{L^p(|x|\geq 1)} \leq C\|V\varphi\|_1 \left(\int_0^\infty |\mathcal{M}\mathcal{H}(M)(\rho)|^p \rho^4 d\rho \right)^{\frac{1}{p}}$$
$$\leq C\|V\varphi\|_1 \left(\int_0^\infty |M(r)|^p r^4 dr \right)^{1/p} \leq C\|V\varphi\|_1^2 \|u\|_p. \tag{5.16}$$

Hölder's inequality and (5.14) for $0 \leq \rho \leq 1$, $K_0^{(j,1)}(\rho) \leq_{|\cdot|} C\mathcal{M}\mathcal{H}(M)(\rho)$, imply

$$|I_1(x)| \leq C \left(\int_{|y|\leq 1} \left| \frac{|V\varphi(x-y)|}{|y|^2} \right|^{p'} dy \right)^{\frac{1}{p'}} \left(\int_0^1 |\mathcal{M}\mathcal{H}(M(r))(\rho)|^p \rho^4 d\rho \right)^{\frac{1}{p}}.$$

Since $p' \leq \frac{5}{4}$ if $p > 5$, Minkowski's inequality and (5.16) imply

$$\|I_1\|_p \leq C\|V\varphi\|_1 \|V\varphi\|_p \|u\|_p. \tag{5.17}$$

(c) We finally prove $-2ia_0C_0(c_1Z_{s0}^{10} - c_0Z_{s0}^{00}) + a_0|\varphi\rangle\langle|D|^{-1}(V\varphi)| \in \mathbf{B}(L^p(\mathbb{R}^5))$ for $p > 5/2$. This will complete the proof of the lemma because this and (b) imply statement (2) by virtue of (5.3); since $|\varphi\rangle\langle|D|^{-1}(V\varphi)|$ is bounded in $L^p(\mathbb{R}^5)$ for $5/3 < p < 5$ as remarked previously, this also implies $-2ia_0C_0(c_1Z_{s0}^{10} - c_0Z_{s0}^{00}) \in \mathbf{B}(L^p(\mathbb{R}^5))$ for $5/3 < p < 5$ and, hence, for $1 < p < 5$ by virtue of result (a) and interpolation. Then, (b) yields statement (1). If $k = 0$, further integration by parts to (5.11) produces boundary term:

$$K_0^{(j,0)}(\rho) = \frac{(-i)^{j+1}}{2\pi} j! \int_{\mathbb{R}} M(r)dr$$
$$+ \frac{(-i)^{j+1}}{2\pi} \int_0^\infty \left(e^{i\lambda\rho}\lambda^j F(\lambda)\right)^{(j+1)} \left(\int_{\mathbb{R}} e^{-i\lambda r} M(r) dr \right) d\lambda. \tag{5.18}$$

The second integral, which we denote by $\tilde{K}_0^{(j,0)}(\rho)$, satisfies

$$\tilde{K}_0^{(j,0)}(\rho) \leq_{|\cdot|} C(1 + \rho^{j+1})\mathcal{M}\mathcal{H}(M)(\rho) \leq C(1 + \rho^{j+2})\mathcal{M}\mathcal{H}(M)(\rho) \tag{5.19}$$

and we estimate the operator $\tilde{Z}^{j0}$ obtained by replacing $K_0^{(j,0)}(\rho)$ by $\tilde{K}_0^{(j,0)}(\rho)$ in (5.4) by repeating the argument of step (b): Split $\tilde{Z}^{j0}u(x)$ as in step (b) and obtain $\|I_2\|_p \leq C\|u\|_p$ for $5/2 < p < 5$ (resp. $p > 5$) by using the first (resp. second) estimate of (5.19) and that r^{4-p} (resp. r^4) is an A_p-weight on $\mathbb{R}$. Likewise we obtain $\|I_1\|_p \leq C\|u\|_p$ for $5/2 < p < 5$ (resp. $p > 5$) by first applying Hölder's inequality by considering the integrand as $(|V\varphi(x-y)|/|y|^2) \cdot (|\tilde{K}_0^{(j,0)}(|y|)|/|y|)$ (resp. $|V\varphi(x-y)|/|y|^3 \cdot |\tilde{K}_0^{(j,0)}(|y|)|$) and then using Minkowski's inequality. Thus, we have for $j = 0, 1$ that

$$\|\tilde{Z}^{j0}u\|_p \leq C\|u\|_p, \quad 5/2 < p < \infty. \tag{5.20}$$

The contribution of boundary terms of (5.18) to $c_0K_0^{(00)} - c_1K_0^{(10)}$ is given by virtue of (2.3) and (3.9) by

$$(c_1 - ic_0) \times \frac{1}{2\pi}\int_{\mathbb{R}} M(r)dr \;=\; \frac{c_0}{\pi i}\int_{\mathbb{R}} M(r)dr \;=\; -4\pi^2 C_0 i\langle |D|^{-1}(V\varphi), u\rangle$$

and this contributes to $2a_0iC_0(c_0Z_{s0}^{00} - c_1Z_{s0}^{10})u(x)$ by

$$8\pi^2 a_0 C_0^2 \int_{\mathbb{R}^5} \frac{V\varphi(y)}{|x-y|^3}dy \cdot (\langle |D|^{-1}(V\varphi), u\rangle) = -a_0\varphi(x)\langle |D|^{-1}(V\varphi), u\rangle,$$

where we used $8\pi^2 C_0 = 1$ when $m = 5$. This proves the lemma. □

5.2 Estimates of Z_{s1} for $m \geq 5$.

We next study $Z_{s1}u$ for all $m \geq 7$. By virtue of (5.6) and (5.7) and the remark at the beginning of section 5, it suffices to study $Z_{1s}^{jk}(\phi)u$ defined by (5.8) for $\phi \in \mathcal{E}$. For simplifying notation, we often omit ϕ from $Z_{1s}^{jk}(\phi)$. Define

$$M_*(r) = M(r, |D|^{-1}(V\phi) * \check{u}). \tag{5.21}$$

Then, by virtue of (2.9), $K^{(j,k)}(\rho)$ may also be expressed as

$$K^{(j,k)}(\rho) = \frac{1}{2\pi}\int_0^\infty e^{i\lambda\rho}\lambda^{j+k}F(\lambda)\left(\int_{\mathbb{R}} e^{-i\lambda r}r^{j+1}M_*(r)dr\right)d\lambda \tag{5.22}$$

which has the larger factor λ^{k+j} than λ^{k+j-1} of (5.9). We omit the proof of the following lemma which is essentially the same as that of (5.13, 5.14)

Lemma 5.2. *$K^{(j,k)}(\rho)$ satisfies the following estimates:*

$$K^{(j,k)}(\rho) \leq_{|\cdot|} \begin{cases} C\rho^{-k-1}\displaystyle\sum_{l=0}^{k+1}\mathcal{M}\mathcal{H}(r^{j+1+l}M)(\rho), & j \geq 2. \quad (5.23) \\ C(1+\rho^{j-1})\mathcal{M}\mathcal{H}(r^2M)(\rho), & j \geq 1. \quad (5.24) \\ C(1+\rho^{j})\mathcal{M}\mathcal{H}(rM)(\rho), & k+j \geq 1. \quad (5.25) \\ C(1+\rho^{j+1})\mathcal{M}\mathcal{H}(M)(\rho), & k \geq 2. \quad (5.26) \\ C(1+\rho^{j})\mathcal{M}\mathcal{H}(rM_*)(\rho), & k \geq 0. \quad (5.27) \end{cases}$$

Lemma 5.3. *Suppose $m \geq 5$ and $\phi \in \mathcal{E}$. Then:*

(1) *If $j \geq 2$, $Z_{1s}^{jk}(\phi)$, $k = 0, \dots, \frac{m-3}{2}$, are bounded in $L^p(\mathbb{R}^m)$ for $1 < p < \frac{m}{2}$.*

(2) *For $k \geq 2$, $Z_{1s}^{jk}(\phi)$, $j = 0, \dots, \frac{m-3}{2}$, are bounded in $L^p(\mathbb{R}^m)$ for $\frac{m}{3} < p$.*

(3) *For all j and k, $Z_{1s}^{jk}(\phi)$ is bounded in $L^p(\mathbb{R}^m)$ for $\frac{m}{3} < p < \frac{m}{2}$.*

If both $j, k \geq 2$, $Z^{jk}_{1s}(\phi)$ is bounded in $L^p(\mathbb{R}^m)$ for all $1 < p < \infty$.

Proof. (a) We first prove (1) for $1 < p < \frac{m}{m-1}$. General case follows from this and (3) by interpolation. We use (5.23) and that $r^{-(m-1)(p-1)}$ is an A_p weight on $\mathbb{R}$ for $1 < p < \frac{m}{m-1}$. Then, estimating as in (5.15), we obtain

$$\begin{aligned}\|Z^{jk}_{s1}u\|_p &\leq C\|V\phi\|_1 \left(\int_0^\infty |M(r)|^p r^{m-1}dr + \int_0^1 \frac{|M(r)|^p}{r^{(m-4)p}} r^{m-1}dr\right)^{1/p} \\ &\leq C\|V\phi\|_1(\|V\phi\|_1 + \|V\phi\|_{p'})\|u\|_p. \end{aligned} \tag{5.28}$$

(b) We next prove (2) for $p > m$. General case then follows from this and (3) by interpolation. We split the integral as in (4.18):

$$Z^{jk}_{s1}u(x) \leq_{|\cdot|} \left(\int_{|y|\leq 1} + \int_{|y|\geq 1}\right) \frac{|V\phi(x-y)|}{|y|^{m-2-k}} |K^{(j,k)}(|y|)|dy = I_1(x) + I_2(x).$$

Using (5.26) for $\rho \geq 1$ and that r^{m-1} is A_p weight on $\mathbb{R}$ if $p > m$, we obtain

$$\|I_2\|_p \leq C\|V\phi\|_1 \left(\int_1^\infty |\mathcal{M}\mathcal{H}(M)(\rho)|^p \rho^{m-1} d\rho\right)^{\frac{1}{p}} \leq C\|V\phi\|_1^2 \|u\|_p. \tag{5.29}$$

Hölder's inequality and (5.26) for $0 \leq \rho \leq 1$ imply that

$$|I_1(x)| \leq \left(\int_{|y|\leq 1} \left|\frac{V\phi(x-y)|}{|y|^{m-2-k}}\right|^{p'} dy\right)^{1/p'} \left(\int_0^1 |\mathcal{M}\mathcal{H}(M)(\rho)|^p \rho^{m-1} d\rho\right)^{1/p}. \tag{5.30}$$

Then, Minkowski's inequality and the estimate as in (5.29) yield

$$\|I_1\|_p \leq C\|V\phi\|_1 \|u\|_p \left(\int_{|x|<1} \frac{\|V\phi\|_p^{p'} dx}{|x|^{(m-2-k)p'}}\right)^{1/p'} \leq C\|V\phi\|_1 \|V\phi\|_p \|u\|_p$$

because $p' \leq \frac{m}{m-1}$ if $p > m$ and $|y|^{-(m-2-k)p'}$ is integrable over $|y| \leq 1$. Thus, statement (2) for $p > m$ follows.

(c) We prove statement (3) by modifying the argument in step (b). Let $\frac{m}{3} < p < \frac{m}{2}$. Then, r^{m-1-2p} is an A_p weight on $\mathbb{R}$. We split the integral of $Z^{jk}_{s1}u(x)$ as in step (b).

(i) Let $j \geq 1$. Estimate (5.24) for $\rho \geq 1$ and Lemma 2.7 yield

$$\|I_2\|_p \leq C\|V\phi\|_1 \left(\int_0^\infty |\mathcal{M}\mathcal{H}(r^2M)(\rho)|^p \rho^{m-1-2p} d\rho\right)^{\frac{1}{p}} \leq C\|V\phi\|_1^2 \|u\|_p. \tag{5.31}$$

Estimate (5.24) for $\rho \leq 1$ and Hölder's inequality imply

$$|I_1(x)| \leq \left(\int_{|y|\leq 1} \left|\frac{|V\phi(x-y)|}{|y|^{m-4-k}}\right|^{p'} dy\right)^{\frac{1}{p'}} \left(\int_0^1 |\mathcal{M}\mathcal{H}(r^2M)(\rho)|^p \rho^{m-1-2p} d\rho\right)^{\frac{1}{p}}.$$

Minkowski's inequality and the second estimate of (5.31) imply $\|I_1\|_p \leq C\|V\phi\|_p\|V\phi\|_1\|u\|_p$ as previously and, hence, $\|Z_{s1}^{jk}u\|_p \leq C\|u\|_p$.
(b) Let $j=0$. Express $K^{(0,k)}(\rho)$ by using $\tilde{M}(r)$ of (2.15) and estimate as

$$K^{(0,k)}(\rho) = \frac{1}{2i\pi}\int_0^\infty e^{i\lambda\rho}\lambda^k\left(\int_{\mathbb{R}} e^{-i\lambda r}\tilde{M}(r)dr\right)F(\lambda)d\lambda \leq_{|\,\cdot\,|} C\mathcal{M}\mathcal{H}(\tilde{M})(\rho). \tag{5.32}$$

Since $\rho^{-(m-2-k)} \leq \rho^{-2}$ for $\rho \geq 1$, Young's inequality, Lemma 2.7 and Hardy's inequality yield

$$\begin{aligned}\|I_2\|_p &\leq C\|V\phi\|_1\left(\int_0^\infty |\tilde{M}(r)|^p r^{m-1-2p}dr\right)^{1/p} \\ &\leq C\|V\phi\|_1\left(\int_0^\infty |M(r)|^p r^{m-1}dr\right)^{1/p} \leq C\|V\phi\|_1\|V\phi\|_p\|u\|_p. \end{aligned}\tag{5.33}$$

Hölder's inequality and (5.32) imply

$$|I_1(x)| \leq \left(\int_{|y|\leq 1}\left|\frac{|V\phi(x-y)|}{|y|^{m-4-k}}\right|^{p'}dy\right)^{1/p'}\left(\int_0^1 |\mathcal{M}\mathcal{H}(\tilde{M})(\rho)|^p\rho^{m-1-2p}d\rho\right)^{1/p}$$

Estimate the second factor by (5.33) and use Minkowski's equality. This yields $\|I_1\|_p \leq C\|V\phi\|_p\|V\phi\|_1\|u\|_p$. The last statement follows from (1) and (2) by interpolation. ☐

Lemma 5.4. *Let $m \geq 5$ and $\phi \in \mathcal{E}$. Then:*

(1) *For $1 < p < \frac{m}{2}$, $\|(c_0 Z_{s1}^{(0,k)} - c_1 Z_{s1}^{(1,k)})u\|_p \leq C\|u\|_p$ for all $0 \leq k \leq \frac{m-3}{2}$.*
(2) *The operator $Z_{s1}(\phi)$ is bounded in $L^p(\mathbb{R}^m)$ for $1 < p < \frac{m}{2}$.*

Proof. It suffices to prove the estimate of (1) for $1 < p < \frac{m}{m-1}$ since that for $1 < p < \frac{m}{2}$ follows from this and Lemma 5.3 (3) by interpolation and since statement (2) follows from this and statement (1) of Lemma 5.3. Using the identity $e^{i\lambda\rho} = (i\rho)^{-k-1}\partial_\lambda^{k+1}e^{i\lambda\rho}$, we apply integration by parts $k+1$ times to the integral of (5.32) and use the identity (2.16). We obtain

$$\begin{aligned}K^{(0,k)}(\rho) &= \frac{i^k}{2\pi\rho^{k+1}}\Bigg(k!\int_{\mathbb{R}} r^2M(r)dr \\ &+\sum_{l=0}^{k+1}\binom{k+1}{l}\int_0^\infty e^{i\lambda\rho}(\lambda^kF)^{(k+1-l)}\int_{\mathbb{R}} e^{-i\lambda r}(-ir)^l\tilde{M}drd\lambda\Bigg).\end{aligned}\tag{5.34}$$

Integration by parts $k+1$ times to $K^{(1,k)}(\rho)$ of (5.9) likewise yields

$$\begin{aligned}K^{(1,k)}(\rho) &= \frac{i^k}{2\pi i\rho^{k+1}}\Bigg(-k!\int_{\mathbb{R}} r^2M(r)dr \\ &-\sum_{l=0}^{k+1}\binom{k+1}{l}\int_0^\infty e^{i\lambda\rho}(\lambda^kF)^{(k+1-l)}\int_{\mathbb{R}} e^{-i\lambda r}(-ir)^l r^2Mdrd\lambda\Bigg).\end{aligned}\tag{5.35}$$

Since $c_0 - ic_1 = 0$, the boundary terms of (5.34) and (5.35) cancel out and

$$\frac{c_0 K^{(0,k)}(\rho) - c_1 K^{(1,k)}(\rho)}{\rho^{m-2-k}} \le_{|\cdot|} \frac{C}{\rho^{m-1}} \sum_{l=0}^{k+1} (\mathcal{MH}(r^l \tilde{M})(\rho) + \mathcal{MH}(r^{l+2}M)(\rho)).$$

For $1 < p < \frac{m}{m-1}$, $\rho^{-(m-1)(p-1)}$ is an A_p-weight on $\mathbb{R}$. It follows by Young's inequality, Lemma 2.7 and Hardy's inequality that $\|(c_0 Z^{(0,k)} - c_1 Z^{(1,k)})u\|_p$ is bounded by $C\|V\phi\|_1$ times

$$\sum_{l=0}^{k+1} \left(\int_0^\infty (|\tilde{M}(r)|^p r^{pl} + |M(r)|^p r^{p(l+2)}) r^{m-1-p(m-1)} dr \right)^{1/p} \tag{5.36}$$

$$\le C \left(\int_0^1 \frac{|M(r)|^p}{r^{p(m-3)}} r^{m-1} dr + \int_0^\infty |M(r)|^p r^{m-1} dr \right)^{1/p} \tag{5.37}$$

$$\le C(\|V\phi\|_{p'} + \|V\phi\|_p)\|u\|_p. \tag{5.38}$$

Here we used $k + 3 \le m - 1$ for $m \ge 5$ in the first step and $p(m-1) < m$ in the last. This proves the estimate of (1) for $1 < p < \frac{m}{m-1}$. □

Lemma 5.1 and the second statement of Lemma 5.4 prove statement (1) of Theorems 1.4 and 1.5 for odd m. The following lemma (and Lemma 5.1 for the case $m = 5$) proves statement (2) of these theorems for odd m.

Lemma 5.5. *Let $m \ge 5$, $\phi \in \mathcal{E}$ and $\frac{m}{2} < p < m$. Then, for a constant $C > 0$,*

$$\left\| Z_{s1}(\phi)u + \frac{\Gamma\left(\frac{m-2}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{m-1}{2}\right)} \langle u, \phi \rangle \phi \right\|_p \le C\|u\|_p. \tag{5.39}$$

If $Z_{s1}(\phi) \in \mathbf{B}(L^p)$ for some $\frac{m}{2} < p < m$, then $\phi \in \mathcal{E}_0$ and $Z_{s1}(\phi) \in \mathbf{B}(L^p)$ for all $1 < p < m$.

Proof. Let $j + k \ge 1$. Since $m - 2 - (k + j) \ge 1$, we have from (5.25) that

$$\frac{K^{(j,k)}(\rho)}{\rho^{m-2-k}} \le_{|\cdot|} C \left(\frac{1}{\rho^{m-2-k}} + \frac{1}{\rho} \right) \mathcal{MH}(rM)(\rho).$$

Using that r^{m-1-p} is A_p weight and $(m-2)p' < m$ for $m/2 < p < m$, we repeat the argument of the step (b) or (c) of the proof of Lemma 5.3 and obtain

$$\|Z_{s1}^{jk} u\|_p \le C\|u\|_p, \quad j + k \ge 1. \tag{5.40}$$

It remains to consider $-2iC_0 c_0 Z_{s1}^{00}$, see (5.7). We apply integration by parts to the right of (5.22) with $j = k = 0$:

$$\begin{aligned} K^{(0,0)}(\rho) &= \frac{i}{2\pi} \int_0^\infty e^{i\lambda\rho} F(\lambda) \partial_\lambda \left(\int_{\mathbb{R}} e^{-i\lambda r} M_*(r) dr \right) d\lambda \\ &= \frac{-i}{2\pi} \int_{\mathbb{R}} M_*(r) dr - \frac{i}{2\pi} \int_0^\infty (e^{i\lambda\rho} F(\lambda))' \left(\int_{\mathbb{R}} e^{-i\lambda r} M_*(r) dr \right) d\lambda. \end{aligned} \tag{5.41}$$

We denote the second integral of (5.41) by $K_*^{(0,0)}(\rho)$ and, by Z_*^{00} the operator produced by substituting $K_*^{(0,0)}(\rho)$ for $K^{(0,0)}(\rho)$ in (5.8). We have $|K_*^{(0,0)}(\rho)| \leq C(1+\rho)\mathcal{M}\mathcal{H}(M_*)(\rho)$. Decompose

$$Z_*^{00}u(x) \leq_{|\,\cdot\,|} \left(\int_{|y|\leq 1} + \int_{|y|\geq 1}\right) |(V\phi)(x-y)| \frac{|K_*^{(0,0)}(|y|)|}{|y|^{m-2}} dy = I_1(x) + I_2(x)$$

as previously. For estimating $\|I_2\|_p$, define $1/q = 1/p - 1/m$ and apply Young's inequality, Hölder's inequality, Lemma 2.7 noticing that $q > m$ and r^{m-1} is A_q weight and, Hardy-Littlewood-Soblev inequality recalling that $|D|^{-1}(V\phi) * (x) \leq_{|\,\cdot\,|} C\langle x\rangle^{1-m}$. We obtain

$$\begin{aligned} \|I_2\|_p &\leq C\|V\phi\|_1 \left(\int_0^\infty |\mathcal{M}\mathcal{H}(M_*)(\rho)|^q \rho^{m-1} d\rho\right)^{1/q} \left\|\frac{1}{|y|^{m-3}}\right\|_{L^m(|y|>1)} \\ &\leq C\|V\phi\|_1 \||D|^{-1}(V\phi) * \check{u}\|_q \leq C\|V\phi\|_1 \||D|^{-1}(V\phi)\|_{\frac{m}{m-1},w}\|u\|_p. \qquad (5.42) \end{aligned}$$

For $I_1(x)$, Hölder's inequality implies

$$|I_1(x)| \leq C\left(\int_{|y|\leq 1}\left|\frac{(V\phi)(x-y)}{|y|^{m-2}}\right|^{q'} dy\right)^{1/q'} \left(\int_{|y|\leq 1} |\mathcal{M}\mathcal{H}(M_*)(|y|)|^q dy\right)^{1/q}.$$

The second factor on the right is bounded by $C\||D|^{-1}(V\phi)\|_{\frac{m}{m-1},w}\|u\|_p$ as in (5.42) and $q' < \frac{m}{m-1} < \frac{m}{2} < p$. It follows by Minkowski's inequality that

$$\|I_1\|_p \leq C\|V\phi\|_p\|u\|_p \left(\int_{|y|\leq 1}\frac{dy}{|y|^{(m-2)q'}}\right)^{1/q'} \leq C\|V\phi\|_p\|u\|_p.$$

Thus, we have $\|Z_*^{00}u\|_p \leq C\|u\|_p$ for $\frac{m}{2} < p < m$. The boundary term of (5.41) is, by virtue of (3.7) and that $c_0 = (m-2)^{-1}$, equal to

$$\begin{aligned} \frac{-i}{2\pi}\int_{\mathbb{R}} M_*(r)dr &= \frac{-i}{\pi\omega_{m-1}}\int_{\mathbb{R}^m}\left(\int_{\mathbb{R}^m}\frac{|D|^{-1}(V\phi)(y)}{|x-y|^{m-1}}dy\right)u(x)dx \\ &= \frac{-i\Gamma\left(\frac{m}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{m-1}{2}\right)}\int_{\mathbb{R}^m}|D|^{-2}(V\phi)(x)u(x)dx = \frac{i\Gamma\left(\frac{m}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{m-1}{2}\right)}\langle\phi, u\rangle. \qquad (5.43) \end{aligned}$$

Inserting this into the right of (5.8) for $j = k = 0$, we see the contribution of the boundary term to $Z_{s1}(\phi)u$ is given by

$$\frac{2c_0C_0\Gamma\left(\frac{m}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{m-1}{2}\right)}\int_{\mathbb{R}^m}\frac{V\phi(y)}{|x-y|^{m-2}}dy\langle\phi, u\rangle = -\frac{\Gamma\left(\frac{m-2}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{m-1}{2}\right)}|\phi\rangle\langle\phi, u\rangle.$$

This proves the first statement. If $Z_{s1} \in \mathbf{B}(L^p)$ for some $\frac{m}{2} < p < m$, (5.39) implies $\phi\otimes\phi \in \mathbf{B}(L^p)$ for this p. Then, (3.8) implies that ϕ must satisfy $\langle\phi, V\rangle = 0$ and $\phi\otimes\phi \in \mathbf{B}(L^p)$ for all $\frac{m}{m-1} < p < m$. Then, $Z_{s1} \in \mathbf{B}(L^p)$ must be satisfied for all $\frac{m}{2} < p < m$ and, hence, for all $1 < p < m$ by Lemma 5.4 and interpolation. $\square$

We finally study $Z_{1s}(\phi)$ in $L^p(\mathbb{R}^m)$ for $p > m$. If $Z_{1s}(\phi) \in \mathbf{B}(L^p(\mathbb{R}^m))$ for some $p > m$, then Lemma 5.5 implies $\phi \in \mathcal{E}_0$. Thus, assume $\phi \in \mathcal{E}_0$ in the following lemma. The following lemma proves statements (3) of Theorem 1.4 and Theorem 1.5 for odd $m \geq 7$.

Lemma 5.6. *Let $m \geq 5$ be odd, $p > m$ and $\phi \in \mathcal{E}_0$. Then:*

1. *For a constant $C_p > 0$, $\|Z_{s1}(\phi)u + |\phi\rangle\langle\phi|)u\|_p \leq C\|u\|_p$.*
2. *If $Z_{1s}(\phi)$ is bounded in $L^p(\mathbb{R}^m)$ for some $p > m$, then $\phi \in \mathcal{E}_1$. In this case Z_{1s} is bounded in $L^p(\mathbb{R}^m)$ for all $1 < p < \infty$.*

Proof. Considering that $\int_{\mathbb{R}} r^{j+1}e^{-i\lambda r}M_*(r)dr = i^{j+1}\left(\int_{\mathbb{R}} e^{-i\lambda r}M_*(r)dr\right)^{(j+1)}$, we apply integration by parts to (5.22). Then, for $k \geq 1$, we have

$$K^{(j,k)}(\rho) = \frac{(-i)^{j+1}}{2\pi}\int_0^\infty \left(e^{i\lambda\rho}\lambda^{j+k}F(\lambda)\right)^{(j+1)}\left(\int_{\mathbb{R}} e^{-i\lambda r}M_*(r)dr\right)d\lambda \quad (5.44)$$

and, if $k = 0$, additional boundary term which is given by virtue of (5.43) by

$$\frac{(-i)^{j+1}j!}{2\pi}\int_{\mathbb{R}} M_*(r)dr = \frac{i(-i)^j j!\Gamma\left(\frac{m}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{m-1}{2}\right)}\langle\phi, u\rangle, \quad j = 0, \dots, \frac{m-3}{2}. \quad (5.45)$$

Denote the right of (5.44) by $\tilde{K}^{(j,0)}(\rho)$ when $k = 0$. Then,

$$\frac{K^{(j,k)}(\rho)}{\rho^{m-2-k}} \leq_{|\cdot|} C\left(1 + \frac{1}{\rho^{m-2}}\right)\mathcal{MH}(M_*)(\rho), \quad 0 \leq j, k \leq \frac{m-3}{2} \quad (5.46)$$

and the same for $\tilde{K}^{(j,0)}(\rho)$. We split $Z_{s1}^{jk}u$ as previously:

$$Z_{s1}^{jk}u(x) = \left(\int_{|x-y|\leq 1} + \int_{|x-y|>1}\right)\frac{V\phi(y)K^{(j,k)}(|x-y|)}{|x-y|^{m-2-k}}dy = I_1(x) + I_2(x).$$

We estimate $I_2(x)$ by using (5.46) for $\rho \geq 1$, that ρ^{m-1} is A_p weight for $p > m$, (3.8) for $\phi \in \mathcal{E}_0$ and the Calderón-Zygmund theory. This yields

$$\begin{aligned}\|I_2\|_p &\leq \|V\phi\|_1\left(\int_1^\infty |\mathcal{MH}(M_*)(\rho)|^p\rho^{m-1}d\rho\right)^{1/p} \\ &\leq \|V\phi\|_1\||D|^{-1}(V\phi) * u\|_p \leq C\|V\phi\|_1\|u\|_p. \quad (5.47)\end{aligned}$$

Hölder's inequality and (5.46) for $\rho \leq 1$ imply

$$|I_1(x)| \leq C\left(\int_{|y|\leq 1}\left|\frac{(V\phi)(x-y)}{|y|^{m-2}}\right|^{p'}dy\right)^{1/p'}\left(\int_{|y|\leq 1}|\mathcal{MH}(M_*)(|y|)|^p dy\right)^{1/p}.$$

The second factor on the right is bounded by $C\|u\|_p$ as in (5.47). Since $p' < \frac{m}{m-1} < m < p$, it follows by Minkowski's inequality that

$$\|I_1\|_p \le C\|V\phi\|_p\|u\|_p \left(\int_{|y|\le 1} \frac{dy}{|y|^{(m-2)p'}}\right)^{1/p'} \le C\|V\phi\|_p\|u\|_p.$$

Thus, $Z_{s1}^{jk} \in \mathbf{B}(L^p(\mathbb{R}^m))$ for $p > m$ if $k \ge 1$ and the same for the operator $\tilde{Z}_{s1}^{j0}$ produced by $\tilde{K}^{(j,0)}(\rho)$. The contribution of boundary terms (5.45) to $Z_{s1}(\phi)$ is given by using the constants C_j of (2.2) by

$$2i\sum_{j=0}^{\frac{m-3}{2}} C_0C_j(-1)^{j+1}\omega_{m-1}\left(\int_{\mathbb{R}^d} \frac{(V\phi)(y)}{|x-y|^{m-2}}dy\right)\frac{i(-i)^j j!\Gamma\left(\frac{m}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{m-1}{2}\right)}\langle\phi, u\rangle$$

$$= -\tilde{D}_m|\phi\rangle\langle\phi, u\rangle, \quad \tilde{D}_m = \sum_{j=0}^{\frac{m-3}{2}} \frac{(m-3-j)!}{2^{m-3-j}\left(\frac{m-3}{2}\right)!\left(\frac{m-3}{2}-j\right)!}. \tag{5.48}$$

The constant $\tilde{D}_m$ can be elementarily computed and with $n = \frac{m-3}{2}$

$$\tilde{D}_m = \sum_{k=0}^{n} \frac{1}{2^{2n-k}}\binom{2n-k}{n-k} = \sum_{k=0}^{n} \frac{1}{2^{n+k}}\binom{n+k}{k} = 1.$$

(see also page 167 of [12].) This proves statement (1). We omit the proof of (2) which is similar to the corresponding statement of Lemma 5.5. ☐

Since $Z_{s1}u = \sum_{i=1}^n Z_{s1}(\phi_j)$ for the orthonormal basis of $\mathcal{E}$, the combination of lemmas in this section proves Theorems 1.4 and 1.5 for odd m.

6 Proof of Theorem 1.5 for even $m \ge 6$

For proving Theorem 1.5 for even dimensions $m \ge 6$ we need study Z_s and $Z_{\log}$ of (3.26) and (3.27). Since $Z_{\log}$ may be studied in a way similar to but simpler than that for Z_s, we shall be mostly concentrated on Z_s and only briefly comment on $Z_{\log}$ at the end of the section. As in odd dimensions we take the real orthonormal basis $\{\phi_1, \ldots, \phi_d\}$ of $\mathcal{E}$ and define, for $\phi \in \mathcal{E}$,

$$Z_s(\phi)u = \frac{i}{\pi}\int_0^\infty G_0(\lambda)|V\phi\rangle\langle\phi V|(G_0(\lambda) - G_0(-\lambda))F(\lambda)\lambda^{-1}d\lambda. \tag{6.1}$$

Then, we have

$$Z_s u = \sum_{j=1}^{d} Z_s(\phi_j)u$$

and we study $Z_s(\phi)$ for $\phi \in \mathcal{E}$. In this section we choose and fix a $\phi \in \mathcal{E}$ arbitrarily and write $M(r) = M(r, V\phi * \check{u})$.

We wish to apply the argument for odd dimensions also to even dimensions as much as possible and, we express $Z_s(\phi)$ as a superposition of operators which are of the same form as those studied in odd dimensions except scaling. We set $\nu = (m-2)/2$. Define for $a > 0$

$$M^a(r) = M(r/(1+2a)) \tag{6.2}$$

and, for $j, k = 0, \ldots, \nu$ and $a, b > 0$,

$$Q^{a,b}_{jk}(\rho) = \frac{(-1)^{j+1}}{2\pi(1+2a)^{j+2}} \int_0^\infty \lambda^{j+k-1} e^{i\lambda(1+2b)\rho} \mathcal{F}(r^{j+1}M^a)(\lambda)F(\lambda)d\lambda. \tag{6.3}$$

As in (5.22), we may express $Q^{a,b}_{jk}(\rho)$ by using $M_*(r)$ and increase the factor λ^{j+k-1} of (6.3) to λ^{j+k}:

$$Q^{a,b}_{jk}(\rho) = \frac{(-1)^{j+1}}{2\pi(1+2a)^{j+2}} \int_0^\infty \lambda^{j+k} e^{i\lambda(1+2b)\rho} \mathcal{F}(r^{j+1}M^a_*)(\lambda)F(\lambda)d\lambda. \tag{6.4}$$

When $j = 0$, we also use $\tilde{M}(r)$ of (2.15) to express $Q^{a,b}_{0k}(\rho)$ as follows:

$$Q^{a,b}_{0k}(\rho) = \frac{i}{2\pi(1+2a)^2} \int_0^\infty \lambda^k e^{i\lambda(1+2b)\rho} \mathcal{F}(\widetilde{M}^a)(\lambda)F(\lambda)d\lambda. \tag{6.5}$$

Lemma 6.1. *Let $Q^{a,b}_{jk}(\rho)$ be defined by* (6.3), (6.4) *or* (6.5). *Then,*

$$Z_s(\phi)u(x) = \frac{2i}{\omega_{m-1}} \sum_{j,k=0}^{\nu} T^{(a)}_j T^{(b)}_k \left[\int_{\mathbb{R}^m} \frac{(V\phi)(x-y)Q^{a,b}_{jk}(|y|)}{|y|^{m-2-k}} dy \right]. \tag{6.6}$$

Proof. We apply (2.18) for $\langle V\phi, (G_0(\lambda) - G_0(-\lambda))u \rangle$ and (2.6) for $G_0(\lambda)$ in (6.1). We see that $Z_s(\phi)u(x)$ is the integral with respect to $\lambda \in (0, \infty)$ of

$$\frac{i}{\pi} \sum_{j,k=0}^{\nu} T^{(a)}_j T^{(b)}_k \left[\frac{(-1)^{j+1}\lambda^{j+k-1}}{(1+2a)^{j+2}\omega_{m-1}} \left(\frac{e^{i\lambda(1+2b)|y|}}{|y|^{m-2-k}} * V\phi \right) \mathcal{F}(r^{j+1}M^a)(\lambda) \right] F(\lambda).$$

Integrating with respect to λ first yields (6.6). □

We define, for $0 \le j, k \le \nu$ and $a, b > 0$, that

$$Z^{jk}(\phi)u(x) = \frac{2i}{\omega_{m-1}} T^{(a)}_j T^{(b)}_k \left[Z^{jk}_{a,b}(\phi)u(x) \right], \tag{6.7}$$

$$Z^{jk}_{a,b}(\phi)u(x) = \int_{\mathbb{R}^m} \frac{(V\phi)(x-y)Q^{a,b}_{jk}(|y|)}{|y|^{m-2-k}} dy. \tag{6.8}$$

Lemma 6.1 implies $Z_s(\phi)u = \sum Z^{jk}(\phi)u$. In what follows we often write $Z^{jk}u$ and $Z^{jk}_{a,b}$ respectively for $Z^{jk}(\phi)u$ and $Z^{jk}_{a,b}(\phi)$.

6.1 Estimate of $\|Z^{jk}u\|_p$ for $(j,k)\neq(\nu,\nu)$.

We estimate Z^{jk} for the case $(j,k)\neq(\nu,\nu)$ first, postponing the case $(j,k)=(\nu,\nu)$ to the next subsection. As we shall see, the argument used for odd dimensions applies to Z^{jk} if $(j,k)\neq(\nu,\nu)$ modulo superpositions and scalings.

Lemma 6.2. *With suitable constants $C>0$, followings are majorants of $Q^{a,b}_{jk}(\rho)$ for $0\le k,j\le\nu$ which satisfy the attached conditions respectively:*

$$(1)\quad C\frac{\{\mathcal{M}\mathcal{H}(r^{j+1}M^a)\}((1+2b)\rho)}{(1+2a)^{j+2}},\quad \textit{if}\ \ j+k\ge 1. \tag{6.9}$$

$$(2)\quad C\frac{\mathcal{M}\mathcal{H}(\widetilde{M}^a)((1+2b)\rho)}{(1+2a)^{2}},\quad \textit{if}\ \ j=0. \tag{6.10}$$

$$(3)\quad C\sum_{l=0}^{k+1}\frac{\mathcal{M}\mathcal{H}(r^{j+l+1}M^a)((1+2b)\rho)}{(1+2a)^{j+2}(1+2b)^{k+1}\rho^{k+1}},\quad \textit{if}\ \ 2\le j\le\nu. \tag{6.11}$$

$$(4)\quad C\frac{\mathcal{M}\mathcal{H}(r^2M^a)((1+2b)\rho)}{(1+2a)^{j+2}}\{(1+2b)^{j-1}\rho^{j-1}+1\},\quad \textit{if}\ \ 1\le j. \tag{6.12}$$

$$(5)\quad C\frac{\mathcal{M}\mathcal{H}(rM^a)((1+2b)\rho)}{(1+2a)^{j+2}}\{(2b+1)^j\rho^j+1\},\quad \textit{for all}\ \ j,k. \tag{6.13}$$

Proof. Define $\Phi_{jk}(\lambda)=\lambda^{j+k-1}F(\lambda)$. If $j+k\ge1$, $\Phi_{jk}\in C_0^\infty(\mathbb{R})$ and Lemma 2.9 implies $Q^{a,b}_{jk}(\rho)=(-1)^{j+1}(1+2a)^{-(j+2)}\{(\mathcal{F}\Phi_{jk})*\mathcal{H}(r^{j+1}M^a)\}((1+2b)\rho)$. Then, (6.9) follows by applying (2.24). Likewise we have (6.10) from (6.5). If $j\ge2$, we apply integration by parts $k+1$ times to (6.3) using that $e^{i\lambda(1+2b)\rho}=(i(1+2b)\rho)^{-(k+1)}\partial_\lambda^{k+1}e^{i\lambda(1+2b)\rho}$ then, without boundary terms,

$$\begin{aligned}Q^{a,b}_{jk}(\rho)&=\sum_{l=0}^{k+1}\frac{(-1)^{j+1}}{2\pi(1+2a)^{j+2}}\left(\frac{1}{-i(1+2b)\rho}\right)^{k+1}\binom{k+1}{l}\\&\times\int_0^\infty e^{i\lambda(1+2b)\rho}\Phi_{jk}(\lambda)^{(k+1-l)}\mathcal{F}((-i)^lr^{j+l+1}M^a)(\lambda)d\lambda\end{aligned} \tag{6.14}$$

and (6.11) follows as previously. If $j\ge1$, we may apply integration by parts to (6.3) by using that $\mathcal{F}(r^{j+1}M^a)(\lambda)=i^{j-1}\{\mathcal{F}(r^2M^a)(\lambda)\}^{(j-1)}$. Then

$$Q^{a,b}_{jk}(\rho)=i^{j-1}\int_0^\infty\frac{\left(\lambda^{j+k-1}F(\lambda)e^{i\lambda(1+2b)\rho}\right)^{(j-1)}\mathcal{F}(r^2M^a)(\lambda)}{2\pi(1+2a)^{j+2}}d\lambda \tag{6.15}$$

and (6.12) follows. Apply another integration by parts in (6.15). No boundary term appears as $\mathcal{F}(rM^a)(0)=0$, and we obtain (6.13). □

6.1.1 Estimate for $1<p<\frac{m}{m-1}$

Define for $0\le\sigma\le m-1$ and $1<p<\frac{m}{m-1}$:

$$N^{a,b}_\sigma(u)=\left(\int_0^\infty|\mathcal{M}\mathcal{H}(r^\sigma M^a)((1+2b)\rho)|^p\rho^{m-1-p(m-1)}d\rho\right)^{1/p}. \tag{6.16}$$

Lemma 6.3. *For any* $\frac{m}{1+\sigma} \leq q \leq \infty$, *we have*

$$N_\sigma^{a,b} \leq C \frac{(1+2b)^{m-1-\frac{m}{p}}}{(1+2a)^{m-1-\frac{m}{p}-\sigma}} (\|V\phi\|_1 + \|V\phi\|_q)\|u\|_p. \tag{6.17}$$

Proof. Change variable ρ by $(1+2b)^{-1}\rho$ first. Since $\rho^{m-1-p(m-1)}$ is an A_p-weight,

$$\begin{aligned} N_\sigma^{a,b} &= (1+2b)^{m-1-\frac{m}{p}} \left(\int_0^\infty |\mathcal{M}\mathcal{H}(r^\sigma M^a)(\rho)|^p \rho^{m-1-p(m-1)} d\rho \right)^{1/p} \\ &\leq C \frac{(1+2b)^{m-1-\frac{m}{p}}}{(1+2a)^{m-1-\frac{m}{p}-\sigma}} \left(\int_0^\infty |M(r)|^p r^{m-1-p(m-1-\sigma)} dr \right)^{1/p}. \end{aligned} \tag{6.18}$$

Denote by I the integral on (6.18). Let $\kappa = m-1-\sigma$. If $\kappa = 0$, then $I \leq C\|V\phi * u\|_p \leq C\|V\phi\|_1\|u\|_p$ and (6.17) follows. Let $0 < \kappa \leq m-1$. Split I into integral over $0 < r < 1$ and $r > 1$ and use $r^{m-1-p\kappa} \leq r^{m-1}$ for $r \geq 1$. Then, we have $I \leq C(\||x|^{-\kappa}(V\phi * u)(x)\|_{L^p(|x|<1)} + \|V\phi\|_1\|u\|_p)$. Take κ' such that $\kappa < \kappa' < m$ and apply Hölder's and Young's inequalities for the integral over $|x| \leq 1$. We obtain with $q = \frac{m}{m-\kappa'} \in \left[\frac{m}{1+\sigma}, \infty\right]$ that

$$I \leq C(\||x|^{-\kappa}\|_{L^{\frac{m}{\kappa'}}(|x|\leq 1)}\|V\phi\|_q + \|V\phi\|_1)\|u\|_p. \tag{6.19}$$

This completes the proof. □

Lemma 6.4. *Suppose* $1 < p < \frac{m}{m-1}$. *Then, for* $2 \leq j \leq \nu$ *and* $0 \leq k \leq \nu$ *such that* $(j,k) \neq (\nu,\nu)$,

$$\|Z^{jk}u\|_p \leq C\|u\|_p, \quad u \in C_0^\infty(\mathbb{R}^m). \tag{6.20}$$

Proof. Minkowski's and Young's inequality imply

$$\|Z^{jk}u\|_p \leq 2\omega_{m-1}^{-1}\|V\phi\|_1 \cdot T_j^{(a)}T_k^{(b)} \left[\left\| |x|^{2+k-m} Q_{jk}^{a,b} \right\|_p \right]. \tag{6.21}$$

We apply (6.11) to estimate $Q_{jk}^{a,b}(|x|)$. Then, since $\sigma \equiv j+l+1 \leq m-1$ for $(j,k) \neq (\nu,\nu)$, Lemma 6.3 implies

$$\left\| |x|^{2+k-m} Q_{jk}^{a,b} \right\|_p \leq C(1+2a)^{\frac{m}{p}-(m-k-1)}(1+2b)^{m-2-\frac{m}{p}-k}\|u\|_p. \tag{6.22}$$

We plug this to (6.21) and use $m-k-1 \geq j+2$. Then,

$$\begin{aligned} \|Z^{jk}u\|_p &\leq C_{mjk}T_j^{(a)}T_k^{(b)}[(1+2a)^{\frac{m}{p}-(j+2)}(1+2b)^{m-2-\frac{m}{p}-k}]\|u\|_p \\ &\leq C\|u\|_p \left(\int_0^\infty \frac{(1+2a)^{\frac{m}{p}-(j+2)}}{(1+a)^{(2\nu-j+\frac{1}{2})}} \frac{da}{\sqrt{a}} \right) \left(\int_0^\infty \frac{(1+2b)^{m-2-\frac{m}{p}-k}}{(1+b)^{(2\nu-k+\frac{1}{2})}} \frac{db}{\sqrt{b}} \right). \end{aligned}$$

Counting powers show that the integrals are finite and the lemma follows. □

As in odd dimensions we use the cancellation in

$$
\begin{aligned}
&Z^{0k}u + Z^{1k}u \\
&\quad = \frac{2i}{\omega_{m-1}} \int_{\mathbb{R}^m} \frac{(V\phi)(x-y)}{|y|^{m-2-k}} T_k^{(b)} (T_0^{(a)} Q_{0k}^{a,b}(|y|) + T_1^{(a)} Q_{1k}^{a,b}(|y|)) dy \qquad (6.23)
\end{aligned}
$$

and obtain the following lemma.

Lemma 6.5. *For $1 < p < \frac{m}{m-1}$, there exists a constant $C > 0$ such that*

$$
\|(Z^{(0,k)} + Z^{(1,k)})u\|_p \le C\|u\|_p, \quad k = 0, \dots, \nu. \tag{6.24}
$$

Proof. We apply integration by parts $k+1$ times to (6.5) and (6.3) as in the proof of (6.11). This produces

$$
Q_{0k}^{a,b}(\rho) = \frac{-i^k k! (\mathcal{F}\widetilde{M^a})(0)\omega_{m-1}}{2\pi(1+2a)^2(1+2b)^{k+1}\rho^{k+1}} - \frac{i^k \omega_{m-1}}{2\pi} \sum_{l=0}^{k+1} C_{k+1,l} Q_{0k,l}^{a,b}(\rho), \tag{6.25}
$$

$$
Q_{1k}^{a,b}(\rho) = \frac{i^{k+1} k! \mathcal{F}(r^2 M^a)(0)\omega_{m-1}}{2\pi(1+2a)^3(1+2b)^{k+1}\rho^{k+1}} + \frac{i^{k+1}\omega_{m-1}}{2\pi} \sum_{l=0}^{k+1} C_{k+1,l} Q_{1k,l}^{a,b}(\rho), \tag{6.26}
$$

where $Q_{0k,l}^{a,b}(\rho)$ and $Q_{1k,l}^{a,b}(\rho)$ are given and estimated as follows:

$$
\begin{aligned}
Q_{0k,l}^{a,b}(\rho) &= \int_0^\infty \frac{e^{i\lambda(1+2b)\rho} (\lambda^k F(\lambda))^{(k+1-l)} (\mathcal{F}((-ir)^l \widetilde{M^a})(\lambda))}{(1+2a)^2(1+2b)^{k+1}\rho^{k+1}} d\lambda \\
&\le_{|\cdot|} C \frac{\mathcal{M}\mathcal{H}(r^l \widetilde{M^a})((1+2b)\rho))}{(1+2a)^2(1+2b)^{k+1}\rho^{k+1}}, \qquad (6.27) \\
Q_{1k,l}^{a,b}(\rho) &= (-i)^l \int_0^\infty \frac{e^{i\lambda(1+2b)\rho} (\lambda^k F(\lambda))^{(k+1-l)} \mathcal{F}(r^{2+l} M^a)(\lambda))}{(1+2a)^3(1+2b)^{k+1}\rho^{k+1}} d\lambda \\
&\le_{|\cdot|} C \frac{\mathcal{M}\mathcal{H}(r^{2+l} M^a)((1+2b)\rho)}{(1+2a)^3(1+2b)^{k+1}\rho^{k+1}}. \qquad (6.28)
\end{aligned}
$$

Eqn.(2.16) shows $\mathcal{F}(\widetilde{M^a})(0) = \mathcal{F}(r^2 M^a)(0) = (1+2a)^3 \int_0^\infty r^2 M(r) dr$ and

$$
T_1^{(a)}[i] = T_0^{(a)}[(1+2a)] = (m-3)^{-1}
$$

It follows that the sum of the superposition via $T_0^{(a)}$ of the boundary term of (6.25) and that via $T_1^{(a)}$ of (6.26) vanishes:

$$
\frac{i^k k!}{(1+2b)^{k+1}\rho^{k+1}} \left(\int_0^\infty r^2 M(r) dr \right) (T_1^{(a)}[i] - T_0^{(a)}[(1+2a)]) = 0. \tag{6.29}
$$

For $1 < p < \frac{m}{m-1}$, $\rho^{m-1-p(m-1)}$ is an A_p weight on $\mathbb{R}$ and we have the identity:

$$
\widetilde{M^a}(r) = \int_r^\infty s M^a(s) ds = (1+2a)^2 \tilde{M}((1+2a)^{-1} r). \tag{6.30}
$$

Then, Lemma 2.7, (6.30), change of variable and Hardy's inequality imply

$$\left\| \frac{Q^{a,b}_{0k,l}(|x|)}{|x|^{m-k-2}} \right\|_p \leq \frac{C(1+2b)^{m-1-\frac{m}{p}}}{(1+2a)^2(1+2b)^{k+1}} \left(\int_0^\infty |r^l \widetilde{M^a}(r)|^p r^{m-1-p(m-1)} dr \right)^{1/p}$$
$$\leq \frac{C(1+2a)^{\frac{m}{p}-(m-1-l)}}{(1+2b)^{\frac{m}{p}-(m-k-2)}} \left(\int_0^\infty |M(r)|^p r^{m-1-p(m-3-l)} dr \right)^{1/p}. \tag{6.31}$$

The integral is similar to the integral which appeared in (6.18) and we remark $m-3-l \geq 0$ for $m \geq 6$. Thus, applying (6.19) with $\sigma = l+2$, we obtain

$$(6.31) \leq \frac{C(1+2a)^{\frac{m}{p}-(m-k-2)}}{(1+2b)^{\frac{m}{p}-(m-k-2)}} (\|V\phi\|_1 + \|V\phi\|_{\frac{m}{3}})\|u\|_p,\ 0 \leq l \leq k+1. \tag{6.32}$$

Counting the powers of a and b, we thus have from (6.32) that

$$T_0^{(a)} T_k^{(b)} \left[\left\| |x|^{2+k-m} Q^{a,b}_{0k,l} \right\|_p \right] \leq C\|u\|_p. \quad 0 \leq l \leq k+1. \tag{6.33}$$

Entirely similarly, starting from (6.28), we obtain

$$\left\| \frac{Q^{a,b}_{1k,l}(|x|)}{|x|^{m-k-2}} \right\|_p \leq \frac{C(1+2b)^{m-1-\frac{m}{p}}}{(1+2a)^3(1+2b)^{k+1}} \left(\int_0^\infty |r^{2+l} M^a(r)|^p r^{m-1-p(m-1)} dr \right)^{1/p}$$
$$\leq \frac{C(1+2a)^{\frac{m}{p}-(m-k-1)}}{(1+2b)^{\frac{m}{p}-(m-k-2)}} (\|V\phi\|_1 + \|V\phi\|_{\frac{m}{3}})\|u\|_p,\ 0 \leq l \leq k+1. \tag{6.34}$$

The extra decaying factor $(1+2a)^{-1}$ of (6.34) compared to (6.32) cancels the extra increasing factor $(1+a)$ of $T_1^{(a)}$ compared to $T_0^{(a)}$ and we have

$$T_1^{(a)} T_k^{(b)} \left[\left\| |x|^{k+2-m} Q^{a,b}_{0k,l}(|x|) \right\|_p \right] \leq C\|u\|_p, \quad 0 \leq l \leq k+1. \tag{6.35}$$

In view of (6.23), (6.25), (6.26) and (6.29), (6.33) and (6.35) with the help of Young's and Minkowski's inequalities imply the lemma. □

6.1.2 Estimate for $\frac{m}{3} < p < \frac{m}{2}$

The following lemma together with Lemma 6.4 and Lemma 6.5 will prove that $\sum_{(j,k)\neq(\nu,\nu)} Z^{jk}$ is bounded in $L^p(\mathbb{R}^m)$ for $1 < p < \frac{m}{2}$.

Lemma 6.6. *Let $\frac{m}{3} < p < \frac{m}{2}$. Then, for $(j,k) \neq (\nu,\nu)$,*

$$\|Z^{jk}u\|_p \leq C_p\|u\|_p \tag{6.36}$$

for a constant $C_p > 0$ independent of $u \in C_0^\infty(\mathbb{R}^m)$.

Proof. Except the superposition the proof is virtually the repetition of that of statement (2) of Lemma 5.3.
(1) Let $j \geq 1$ first. Since ρ^{m-1-2p} is A_p weight for $\frac{m}{3} < p < \frac{m}{2}$, we have

$$\left(\int_0^\infty |\{\mathcal{M}\mathcal{H}(r^2 M^a)\}(\rho)|^p \rho^{m-1-2p} d\rho\right)^{\frac{1}{p}} \leq C(1+2a)^{\frac{m}{p}} \|V\phi\|_1 \|u\|_p. \quad (6.37)$$

Splitting the integral of (6.8) we define

$$Z_{jk}^{a,b} u(x) = \left(\int_{|y|<\frac{1}{1+2b}} + \int_{|y|>\frac{1}{1+2b}}\right) \frac{(V\phi)(x-y) Q_{jk}^{a,b}(|y|)}{|y|^{m-2-k}} dy = I_1(x) + I_2(x). \quad (6.38)$$

For $I_1(x)$, we estimate $|y|^{-(m-k-2)} \leq |y|^{-(m-2)}$ for $|y| \leq 1$ and apply Hölder's inequality. Then

$$\|I_1\|_p \leq \left\| \int_{|y|\leq \frac{1}{2b+1}} \frac{|(V\phi)(x-y)|^{p'} dy}{|y|^{p'(m-4)}} \right\|_{p/p'}^{1/p'} \left(\int_{|y|\leq \frac{1}{2b+1}} \left|\frac{Q_{jk}^{a,b}(|y|)}{|y|^2}\right|^p dy\right)^{1/p}$$

Minkowski's inequality implies that the first factor on the right is bounded by $C\|V\phi\|_p (1+2b)^{m-4-\frac{m}{p'}}$ and $\frac{m}{p'} - (m-4) > 1$. For the second factor, we apply (6.12) for $(1+2b)\rho < 1$ and then (6.37). We obtain

$$\|I_1\|_p \leq C(1+2a)^{\frac{m}{p}-j-2} (1+2b)^{1-\frac{m}{p}} \|V\phi\|_1 \|V\phi\|_p \|u\|_p. \quad (6.39)$$

By Young's inequality $\|I_2\|_p \leq C\|V\phi\|_1 \||x|^{2+k-m} Q_{jk}^{a,b}(|x|)\|_{L^p((1+2b)|x|>1)}$. For the second factor, we use (6.12) for $(1+2b)\rho \geq 1$ and, after changing the variables $\rho \to (1+2b)^{-1}\rho$, we estimate $\rho^{-(m-2-k-(j-1))} \leq \rho^{-2}$ for $\rho \geq 1$ (here we used $(j,k) \neq (\nu,\nu)$) and apply (6.37) once more. Then,

$$\begin{aligned} \|I_2\|_p &\leq C\|V\phi\|_1 \frac{(1+2b)^{m-2-k-\frac{m}{p}}}{(1+2a)^{j+2}} \left(\int_1^\infty |\{\mathcal{M}\mathcal{H}(r^2 M^a)\}(\rho)|^p \rho^{m-1-2p} d\rho\right)^{\frac{1}{p}} \\ &\leq C(1+2a)^{\frac{m}{p}-j-2} (1+2b)^{m-2-k-\frac{m}{p}} \|V\phi\|_1^2 \|u\|_p. \end{aligned} \quad (6.40)$$

Since $m-2-k \geq 1$ and $(1+2a)^{\frac{m}{p}-j-2}(1+2b)^{m-2-k-\frac{m}{p}}$ is summable by $T_j^{(a)} T_k^{(b)}$, (6.39) and (6.40) imply

$$\|Z^{jk} u\|_p \leq C\|V\phi\|_1 (\|V\phi\|_1 + \|V\phi\|_p) \|u\|_p. \quad (6.41)$$

(2) When $j = 0$, we apply the argument in the proof in (1) for estimating $Q_{0k}^{a,b}$ but by using (6.10) in stead of (6.12). Then, by the help of (6.30) and Hardy's inequality, it leads to estimates (6.39) and (6.40) and, hence, to the desired (6.36) for Z^{0k}. This completes the proof of the lemma. ∎

6.1.3 Estimate for $m/2 < p < m$ and for $p > m$.

We now estimate Z^{jk}, $(j,k) \neq (\nu,\nu)$, in $L^p(\mathbb{R}^m)$ for $\frac{m}{2} < p < m$ and for $p > m$. As in odd dimensions, Z^{00} will not in general be bounded in $L^p(\mathbb{R}^m)$ when $\frac{m}{2} < p < m$ and likewise for all Z^{0k}, $k = 0, \dots, \frac{m-2}{2}$ when $p > m$. Elementary computations using

$$z^{-k} = \frac{1}{\Gamma(k)} \int_0^\infty e^{-zt} t^{k-1} dt, \quad \Re z > 0, \quad k > 0$$

and the formula (2.5) for $C_{m,j}\omega_{m-1}$ we obtain the following lemma.

Lemma 6.7.

(1) *We have* $T_j^{(a)}[1] = (m-3-j)!/(m-2)!$.

(2) *For* $k \geq 1$ *and* $j = 0, \cdots, \nu$, $T_j^{(a)}[(1+2a)^{-k}]$ *is given by*

$$\frac{(-i)^j 2^{m-1}\Gamma(2\nu - j + k)}{(m-2)!\Gamma(k)} \binom{\nu}{j} \cdot \int_1^\infty \frac{(x^2-1)^{k-1}}{(x^2+1)^{2\nu-j+k}} dx \tag{6.42}$$

Lemma 6.8. *Let* $\frac{m}{2} < p < m$ *and* $\phi \in \mathcal{E}$. *Then:*

(1) *If* $(j,k) \neq (0,0)$ *or* $(j,k) \neq (\nu,\nu)$, Z^{jk} *is bounded in* $L^p(\mathbb{R}^m)$:

$$\|Z^{jk}u\|_p \leq C\|u\|_p, \quad u \in C_0^\infty(\mathbb{R}^m) \tag{6.43}$$

(2) *There exists a constant* $C > 0$ *such that for* $u \in C_0^\infty(\mathbb{R}^m)$, *we have*

$$\left\| Z^{00}u + D_m|\phi\rangle\langle\phi, u\rangle \right\|_p \leq C\|u\|_p, \tag{6.44}$$

$$D_m = \frac{2^m \Gamma\left(\frac{m}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{m-1}{2}\right)} \int_1^\infty (1+x^2)^{m-1} dx. \tag{6.45}$$

If $Z^{00}(\phi)$ *is bounded in* $L^p(\mathbb{R}^m)$ *for some* $\frac{m}{2} < p < m$ *then* $\phi \in \mathcal{E}_0$. *In this case* $Z^{00}(\phi)$ *is bounded in* $L^p(\mathbb{R}^m)$ *for all* $\frac{m}{2} < p < m$.

Proof. (1) Split $Z^{jk}_{a,b}u(x)$ as in (6.38) and apply the argument thereafter to $I_1(x)$ and $I_2(x)$ by using the estimate (6.13). Since $m-2-(k+j) \geq 1$ and ρ^{m-1-p} is an A_p weight for $\frac{m}{2} < p < m$, we have, as in (6.40),

$$\|I_2\|_p \leq C\frac{(1+2b)^{m-2-k-\frac{m}{p}}}{(1+2a)^{j+2-\frac{m}{p}}} \|V\phi\|_1^2 \|u\|_p. \tag{6.46}$$

For dealing with $I_1(x)$, we estimate $|y|^{-(m-2-k)} \leq |y|^{m-2}$ for $|y| \leq 1$ as previously but now decompose $|y|^{-(m-2)} = |y|^{-(m-3)} \cdot |y|^{-1}$, remarking that $(m-3)p' < m$ and $p/p' > 1$. Then, we obtain as in (6.39) that

$$\|I_1\|_p \leq \frac{(1+2a)^{\frac{m}{p}-j-2}}{(1+2b)^{\frac{m}{p}}} \|V\phi\|_1 \|V\phi\|_p \|u\|_p. \tag{6.47}$$

Summing up (6.46) and (6.47) by $T_j^{(a)}T_k^{(b)}$, we obtain (6.43).
(2) Let $j = k = 0$. We apply integration by parts to (6.4).

$$Q_{00}^{a,b}(\rho) = \frac{-i}{2\pi(1+2a)^2}\int_0^\infty e^{i\lambda(1+2b)\rho}\mathcal{F}(M_*^a)'(\lambda)F(\lambda)d\lambda$$
$$= \frac{i}{2\pi}\int_{\mathbb{R}}\frac{M_*^a(r)}{(1+2a)^2}dr + \frac{i}{(1+2a)^2}\int_0^\infty (F(\lambda)e^{i\lambda(1+2b)\rho})'\mathcal{F}(M_*^a)(\lambda)d\lambda. \quad (6.48)$$

Denote the second term on (6.48) by $\tilde{Q}_{00}^{ab}(\rho)$ and by $\tilde{Z}^{00}$ the operator produced by inserting $\tilde{Q}_{00}^{ab}(\rho)$ for $Q_{00}^{ab}(\rho)$ in (6.7). We have

$$\tilde{Q}_{00}^{a,b}(\rho) \le_{|\,\cdot\,|} C\frac{\mathcal{MH}(M_*^a)((1+2b)\rho)}{(1+2a)^2}(1+(1+2b)\rho). \quad (6.49)$$

Let $\frac{m}{2} < p < m$. We split as in (6.38) and estimate I_2 first:

$$\tilde{Z}^{00}u(x) = \left(\int_{|y|<\frac{1}{1+2b}} + \int_{|y|\ge\frac{1}{1+2b}}\right)\frac{(V\phi)(x-y)\tilde{Q}_{00}^{a,b}(|y|)}{|y|^{m-2}}dy = I_1(x) + I_2(x).$$

We obtain

$$\begin{aligned}\|I_2\|_p &\le C\|V\phi\|_1\frac{(1+2b)^{m-2-\frac{m}{p}}}{(1+2a)^2}\left\|\frac{\mathcal{MH}(M_*^a)(|y|)}{|y|^{m-3}}\right\|_{L^p(|y|>1)} \\ &\le C\|V\phi\|_1\frac{(1+2b)^{m-2-\frac{m}{p}}}{(1+2a)^2}\left\|\frac{1}{|y|^{m-3}}\right\|_{L^m(|y|>1)}\left(\int_0^\infty |M_*^a(r)|^q r^{m-1}dr\right)^{\frac{1}{q}} \\ &\le C\frac{\|V\phi\|_1(1+2b)^{m-2-\frac{m}{p}}}{(1+2a)^{2-\frac{m}{q}}}\||D|^{-1}(V\phi)\|_{\frac{m}{m-1},\infty}\|u\|_p, \end{aligned} \quad (6.50)$$

where we used Young's inequality, (6.49) for $(1+2b)\rho \ge 1$ and the change of variable $(1+2b)\rho$ to ρ in the first step, Hölder's inequality considering $p^{-1} = m^{-1} + q^{-1}$ and that 1 is an A_q weight $q = mp/(m-p) > m$ in the second and finally weak-Young's inequality. For I_1, we apply Hölder's and Minkowski's inequalities and (6.50) and obtain

$$\begin{aligned}\|I_1\|_p &\le C\left\|\left(\int_{|y|\le\frac{1}{1+2b}}\left|\frac{(V\phi)(x-y)}{|y|^{m-2}}\right|^{q'}dy\right)^{\frac{1}{q'}}\right\|_p \\ &\quad\times(1+2b)^{-\frac{m}{p}}(1+2a)^{-2}\left(\int_{|y|\le 1}|\mathcal{MH}(M_*^a)(|y|)|^q dy\right)^{1/q} \\ &\le C(1+2b)^{-\frac{m}{p}}(1+2a)^{\frac{m}{p}-2}\|V\phi\|_p\||D|^{-1}(V\phi)\|_{\frac{m}{m-1},\infty}\|u\|_p. \end{aligned} \quad (6.51)$$

Summing (6.50) and (6.51) by $T_0^{(a)}T_0^{(b)}$, we obtain $\|\tilde{Z}^{(0,0)}u\|_p \le C\|u\|_p$.

By virtue of (3.10) and (5.43), the contribution to $Z^{00}u$ of the boundary term of (6.48) is given by

$$\frac{2i}{\omega_{m-1}}T_0^{(a)}T_0^{(b)}\left[\int_{\mathbb{R}^m}\frac{(V\phi)(y)dy}{|x-y|^{m-2}}\cdot\frac{i}{2\pi}\int_{\mathbb{R}}\frac{M_*(r)}{(1+2a)}dr\right]$$
$$=-\frac{2}{C_0\omega_{m-1}}T_0^{(a)}[(1+2a)^{-1}]T_0^{(b)}[1]\frac{\Gamma\left(\frac{m}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{m-1}{2}\right)}\langle\phi,u\rangle\phi. \qquad (6.52)$$

By using Lemma 6.7 and $C_0\omega_{m-1}=(m-2)^{-1}$. we can simplify (6.52) to $-D_m\langle\phi,u\rangle\phi$ with D_m given by (6.45) and (6.44) follows. The last statement follows as in the odd dimensional case, see the remark after Lemma 5.5. □

Finally in this section we study $Z^{jk}(\phi)u$ for $(j,k)\neq(\nu,\nu)$ in $L^p(\mathbb{R}^m)$ when $p>m$, assuming $\phi\in\mathcal{E}_0$ by the same reason as in odd dimensions. We define

$$D_{m,j}=2^m\binom{\nu}{j}\frac{\Gamma\left(\frac{m}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{m-1}{2}\right)}\int_1^\infty\frac{(x^2-1)^j}{(x^2+1)^{m-1}}dx,\quad j=0,\dots,\nu. \qquad (6.53)$$

Lemma 6.9. *Let $m\geq 6$ be even and $p>m$. Suppose that $\phi\in\mathcal{E}_0$. Then:*

(1) *For (j,k) such that $k\neq 0$ and $(j,k)\neq(\nu,\nu)$, Z^{jk} is bounded in $L^p(\mathbb{R}^m)$.*

(2) *There exists a constant $C>0$ such that*

$$\|Z^{j0}u+D_{j,m}\langle\phi,u\rangle\phi\|_p\leq C\|u\|_p,\quad j=0,\dots,\nu. \qquad (6.54)$$

(3) *If $Z^{j0}(\phi)$ is bounded in $L^p(\mathbb{R}^m)$ for some $0\leq j\leq\nu$ and some $m<p<\infty$, then $\phi\in\mathcal{E}_1$. In this case, $Z^{j0}(\phi)$ is bounded in $L^p(\mathbb{R}^m)$ for all $1<p<\infty$ and $0\leq j\leq\nu$.*

Proof. We apply integration by parts $j+1$ times to (6.4):

$$Q_{jk}^{a,b}(\rho)=\int_0^\infty\frac{(-i)^{j+1}\lambda^{j+k}F(\lambda)e^{i\lambda(1+2b)\rho}\partial_\lambda^{j+1}\{\mathcal{F}(M_*^a)(\lambda)\}}{2\pi(1+2a)^{j+2}}d\lambda. \qquad (6.55)$$

(1) If $k\geq 1$, then no boundary terms appear and we have

$$Q_{jk}^{a,b}(\rho)\leq_{|\cdot|}\frac{C\mathcal{M}\mathcal{H}(M_*^a)((1+2b)\rho)}{(1+2a)^{j+2}}\{(1+2b)^{j+1}\rho^{j+1}+1\}. \qquad (6.56)$$

Observing that $m-2-(k+j+1)\geq 0$ for $(j,k)\neq(\nu,\nu)$, that r^{m-1} is A_p weight on $\mathbb{R}$ for $p>m$ and that $(m-2-k)p'<m$, we apply the argument used for proving (6.50) and (6.51) in the proof of the previous lemma and obtain

$$\|Z_{a,b}^{jk}u\|_p\leq\frac{C\|V\phi\|_1(1+2b)^{m-2-k-\frac{m}{p}}}{(1+2a)^{j+2-\frac{m}{p}}}\||D|^{-1}(V\phi)*u\|_p$$
$$+\frac{C(1+2b)^{-\frac{m}{p}}\|V\phi\|_p}{(1+2a)^{j+2-\frac{m}{p}}}\left(\int_{|y|<\frac{1}{1+2b}}\frac{dy}{|y|^{(m-2-k)p'}}\right)^{\frac{1}{p'}}\||D|^{-1}(V\phi)*u\|_p. \qquad (6.57)$$

Since $\int (V\phi)(x)dx = 0$, $\| |D|^{-1}(V\phi) * u\|_p \le C\|u\|_p$ for any $1 < p < \infty$ by virtue of (3.9) and the Calderón-Zygmund theory. It follows that

$$\|Z^{jk}u\|_p \le T_j^{(a)} T_k^{(b)} \|Z^{jk}_{a,b}u\|_p \le C\|u\|_p$$

for $k \ge 1$ and $(j,k) \ne (\nu,\nu)$ and statement (1) is proved.
(2) If $k = 0$, then, $j+1$ times integration by parts in (6.55) produces the integral and boundary terms. The integral is bounded by (6.56) and the repetition of the argument in step (1) implies that its contribution to Z^{j0} is the operator which is bounded in $L^p(\mathbb{R}^m)$ for all $p > m$. The boundary term may be expressed as follows by using (5.43) once more,

$$\frac{i^{j+1}j!}{2\pi(1+2a)^{j+1}} \int_{\mathbb{R}} M_*(r)dr = \frac{-i^{j+1}j!}{(1+2a)^{j+1}} \frac{\Gamma\left(\frac{m}{2}\right)}{\Gamma\left(\frac{m-1}{2}\right)\sqrt{\pi}} \langle \phi, u\rangle, \tag{6.58}$$

and its contribution $Z^{j0}u$ may be computed as follows:

$$\begin{aligned} &\frac{2i}{\omega_{m-1}} T_j^{(a)} T_0^{(b)} \left[\int_{\mathbb{R}^m} \frac{(V\phi)(y)dy}{|x-y|^{m-2}} \left(-\frac{i^{j+1}j!}{(1+2a)^{j+1}}\right)\right] \frac{\Gamma\left(\frac{m}{2}\right)}{\Gamma\left(\frac{m-1}{2}\right)\sqrt{\pi}} \langle \phi, u\rangle \\ &\qquad = 2i^{j+2}j! T_j^{(a)}[(1+2a)^{-(j+1)}] \frac{\Gamma\left(\frac{m}{2}\right)}{\Gamma\left(\frac{m-1}{2}\right)\sqrt{\pi}} \langle \phi, u\rangle\phi = -D_{m,j}\langle\phi, u\rangle\phi. \end{aligned}$$

where we used $C_0\omega_{m-1} = T_0^{(b)}[1] = (m-2)^{-1}$ and (6.42) with $k = j+1$ for $T_j^{(a)}[(1+2a)^{-(j+1)}]$. This proves statement (2). We omit the proof of statement (3) which is similar to the corresponding part of the previous lemma. □

Lemma 6.10. *Define* $\tilde{D}_m = \sum_{j=0}^{\nu} D_{m,j}$. *Then,* $\tilde{D}_m = 1$.

Proof. Use binomial formula for (6.53). We have

$$\tilde{D}_m = 2^m \frac{\Gamma\left(\frac{m}{2}\right)}{\Gamma\left(\frac{m-1}{2}\right)\sqrt{\pi}} \int_1^\infty \frac{x^{m-2}}{(x^2+1)^{m-1}} dx$$

Change of variable $x \to x^{-1}$ shows that the integral is equal to the same integral over the interval $0 < x < 1$. It follows after making the change of variable $x^2 = t$ that the integral is equal to

$$\frac{1}{4}\int_0^\infty \frac{t^{\nu-\frac{1}{2}}}{(t+1)^{m-1}} dt = \frac{\Gamma\left(\frac{m-1}{2}\right)^2}{2^2\Gamma(m-1)}.$$

Thus, $\tilde{D}_m = 2^{m-2}\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{m-1}{2}\right)\Gamma(m-1)^{-1}\pi^{-\frac{1}{2}} = 1$. □

In the next two sections we prove that $Z^{\nu\nu}$ and $Z_{\log}$ are bounded in $L^p(\mathbb{R}^m)$ for all $1 < p < \infty$. These will complete the proof of Theorem 1.5.

6.2 Estimate of $\|Z^{\nu\nu}u\|_p$ for $1<p<\infty$

In this section we prove

$$\|Z^{\nu\nu}u\|_p \le C\|u\|_p, \quad 1<p<\infty. \tag{6.59}$$

The method of previous subsection does not apply for proving this and we exploit more direct method. By virtue of interpolation, it suffices to prove (6.59) for arbitrarily small $p>1$ and large $p>m$.

6.2.1 The case for $1<p<\frac{2(m-1)}{m+1}$

We first show (6.59) for $1<p<\frac{2(m-1)}{m+1}$. After changing the variable r to $(1+2a)r$ in (6.3), we write $Q^{a,b}_{\nu\nu}(\rho)/\rho^\nu$ in the form

$$\frac{(-1)^{\nu+1}}{2\pi\rho^\nu}\int_0^\infty e^{i(1+2b)\rho\lambda}\lambda^{m-3}F(\lambda)\left(\int_{\mathbb{R}} e^{-i(1+2a)r\lambda}r^{\nu+1}M(r)dr\right)d\lambda. \tag{6.60}$$

Integration by parts implies that (6.60) is equal to

$$\begin{aligned}&\frac{i(-1)^{\nu+1}}{2\pi(1+2b)\rho^{\nu+1}}\int_0^\infty e^{i(1+2b)\rho\lambda}(\lambda^{m-3}F(\lambda))'\left(\int_{\mathbb{R}} e^{-i(1+2a)r\lambda}r^{\nu+1}M(r)dr\right)d\lambda\\ &+\frac{(-1)^{\nu+1}(1+2a)}{2\pi(1+2b)\rho^{\nu+1}}\int_0^\infty e^{i(1+2b)\rho\lambda}\lambda^{m-3}F(\lambda)\left(\int_{\mathbb{R}} e^{-i(1+2a)r\lambda}r^{\nu+2}Mdr\right)d\lambda.\end{aligned}$$

The first line becomes $i(1+2b)^{-1}Q^{a,b}_{\nu(\nu-1)}(\rho)/\rho^{m-2-(\nu-1)}$ if we replace $(m-3)F(\lambda)+\lambda F'(\lambda)$ by $F(\lambda)$ and the former function can play the same role as the latter does in the argument of previous sections and, $\nu-1\ge 1$ if $m\ge 6$. Thus, if we substutute it for $Q^{a,b}_{\nu\nu}(\rho)/\rho^\nu$ in (6.8) for $(j,k)=(\nu,\nu)$ and, then the resulting function for $Z^{\nu\nu}_{a,b}(\phi)u(x)$ in (6.7), it produces the operator which has the same L^p property as $Z^{\nu(\nu-1)}$ which is bounded in $L^p(\mathbb{R}^m)$ for $1<p<\infty$. Hence, we need study only the operator produced by the second line. Once again we substitute it for $Q^{a,b}_{\nu\nu}(\rho)/\rho^\nu$ in (6.8) and the result for $Z^{\nu\nu}_{a,b}(\phi)u(x)$ in (6.7). We denote the functiotn thus obtain by $Z^{\nu\nu}u(x)$, abusing notation. We want to show that this $Z^{\nu\nu}u(x)$ satisfies (6.59) for $1<p<\frac{m}{m-1}$. Integrating with respect to a,b first via Fubini's theorem shows

$$Z^{\nu\nu}u(x)=\frac{2i}{\omega_{m-1}}\int_{\mathbb{R}^m}(V\phi)(x-y)X_\nu(|y|)dy, \tag{6.61}$$

$$\begin{aligned}X_\nu(\rho)=\frac{2iC^2_{m,\nu}\omega_{m-1}}{\rho^{\nu+1}}&\int_0^\infty\left\{e^{i\lambda\rho}\lambda^{m-3}\left(\int_0^\infty\frac{(1+2b)^{-1}e^{2i\lambda\rho b}}{(1+b)^{\nu+\frac12}}\frac{db}{\sqrt b}\right)\right.\\ &\times\left.\int_{\mathbb{R}}e^{-i\lambda r}\left(\int_0^\infty\frac{(1+2a)e^{-2ia\lambda r}}{(1+a)^{\nu+\frac12}}\frac{da}{\sqrt a}\right)r^{\nu+2}M(r)dr\right\}F(\lambda)d\lambda.\end{aligned} \tag{6.62}$$

Let $\chi_\pm(r) = 1$ for $\pm r > 0$ and $\chi_\pm(r) = 0$ for $\pm r \le 0$. Define, for $t > 0$,

$$g_\pm(t) = \int_0^\infty \left(1 + \frac{a}{t}\right)\left(1 + \frac{a}{2t}\right)^{-\nu - \frac{1}{2}} e^{\pm ia} \frac{da}{\sqrt{a}}, \tag{6.63}$$

$$h_\pm(t) = \int_0^\infty \left(1 + \frac{b}{t}\right)^{-1}\left(1 + \frac{b}{2t}\right)^{-\nu - \frac{1}{2}} e^{\pm ib} \frac{db}{\sqrt{b}} \tag{6.64}$$

and, with $C = iC_{m,\nu}^2 \omega_{m-1}$, write $X_\nu(\rho)$ as follows:

$$X_\nu(\rho) = \frac{C}{\rho^{\nu + \frac{3}{2}}} \int_{\mathbb{R}} (L_+(\rho, r) + L_-(\rho, r)) r^{\nu+2} |r|^{-\frac{1}{2}} M(r) dr, \tag{6.65}$$

$$L_\pm(\rho, r) = \chi_\pm(r) \int_0^\infty e^{i\lambda(\rho - r)} \lambda^{m-4} h_+(\lambda\rho) g_\mp(\pm r\lambda) F(\lambda) d\lambda. \tag{6.66}$$

Lemma 6.11. *Suppose that f is C^∞ on $[0,\infty)$ and satisfies $|f^{(j)}(c)| \le C_j c^{-(j+1)}$ for $c \ge 1$, $j = 0, 1, \dots$. Define*

$$\ell_\pm(t) = \int_0^\infty e^{\pm ic} f(c/t) \frac{dc}{\sqrt{c}}.$$

Then, $\ell_\pm(t)$ is C^∞ for $t > 0$ and satisfies the following properties.

(1) *$\ell_\pm(1/t)$ can be exteded to a C^∞ function on $[0, 1]$, hence, $\lim_{t\to\infty} \ell_\pm(t) = \alpha_\pm$ exists and for $t \ge 1$, $|\ell_\pm^{(j)}(t)| \le C_j t^{-j-1}$, $j = 1, 2, \dots$.*

(2) *For $0 < t < 1$, $|t^j \ell_\pm^{(j)}(t)| \le C_j \sqrt{t} \le C_j$, $j = 0, 1, \dots$.*

Proof. We prove the lemma for $\ell_+(t)$ only and omit the $+$-sign. It is evident that $\ell(t)$ is C^∞ for $t > 0$. Splitting the interval, we define

$$\ell(t) = \left(\int_0^1 + \int_1^\infty\right) f\left(\frac{c}{t}\right) e^{ic} \frac{dc}{\sqrt{c}} \equiv \ell_1(t) + \ell_2(t).$$

It is obvious that $\ell_1(1/t)$ is of $C^\infty[0, 1]$. To see the same for $\ell_2(1/t)$, we perform integration by parts n times for $t > 0$:

$$i^n \ell_2(1/t) = B_n(t) + (-1)^n \int_1^\infty \partial_c^n \left(\frac{f(ct)}{\sqrt{c}}\right) e^{ic} dc. \tag{6.67}$$

The boundary term $B_n(t)$ is a polynomial of order n and Leibniz' formula implies $\partial_c^n \left(\frac{f(ct)}{\sqrt{c}}\right) = \sum_{j=0}^n C_{nj} f^{(j)}(ct)(ct)^j c^{-\frac{1}{2} - n}$. Since $\partial_y^k (f^{(j)}(y) y^j)$ is bounded for any $j, k = 0, 1, \dots$ and

$$\partial_t^k \left(\sum_{j=0}^n C_{nj} f^{(j)}(ct)(ct)^j c^{-\frac{1}{2} - n}\right) = \sum_{j=0}^n C_{nj}\ \partial_y^k (f^{(j)}(y) y^j)\Big|_{y=ct} c^{-\frac{1}{2} - n + k},$$

the integral of (6.67) is a function of class $C^{n-1}([0,1])$. Since n is arbitray, this proves (1). For proving (2), after changing the variable we decompose:

$$\ell(t) = \sqrt{t}\left(\int_0^1 + \int_1^\infty\right) f(c)e^{ict}\frac{dc}{\sqrt{c}} \equiv \sqrt{t}(\tilde{\ell}_1(t) + \tilde{\ell}_2(t))$$

We obseve that $\sqrt{t}$ satisfies the property (2) and that, if $\alpha(t)$ satisfies (2) and $|t^j\beta^{(j)}(t)| \le C_j$, then so does $\gamma(t) = \alpha(t)\beta(t)$. Hence, $\sqrt{t}\tilde{\ell}_1(t)$ satisfies (2) because $\tilde{\ell}_1(t)$ is entire. To prove the same for $\sqrt{t}(\tilde{\ell}_2(t)$, it suffices to show that $|(t^n\tilde{\ell}_2(t))^{(n)}| \le C_n$ for $0 < t < 1$, $n = 0, 1, 2, \dots$. By integration by parts we have

$$\begin{aligned}(it)^n\tilde{\ell}_2(t) &= \int_1^\infty (\partial_c^n e^{itc}) f(c)\frac{dc}{\sqrt{c}} \\ &= \sum_{j=0}^{n-1}(-1)^{j+1}\partial_c^j\left(\frac{f(c)}{\sqrt{c}}\right)\partial_c^{n-j-1}(e^{itc})\big|_{c=1} + \int_1^\infty e^{itc}(f(c)c^{-\frac{1}{2}})^{(n)}dc.\end{aligned}$$

The boundary term is a polynomial of t and the integral is n times continuously differentiable and a fortiori $(t^n\tilde{\ell}_2(t))^{(n)} \le C$ for $0 < t < 1$. □

We define $L_{\pm,\sigma}(\rho, r)$ for an integer $\sigma \ge 0$ and functions $g_\pm$ and h by

$$L_{\pm,\sigma}(\rho, r) = \chi_\pm(r)\int_0^\infty e^{i\lambda(\rho-r)}\lambda^\sigma h_+(\lambda\rho)g_\mp(\pm r\lambda)F(\lambda)d\lambda \tag{6.68}$$

so that we have $L_\pm(\rho, r) = L_{\pm,m-4}(\rho, r)$ (see (6.66)).

Lemma 6.12. *Suppose that $g_\pm(t)$ and $h_+(t)$ are C^∞ functions of $t > 0$ and they satisfy following properties replacing f:*

(a) *The limit* $\lim_{t\to\infty} f(t)$ *exists.*

(b) $|t^j f^{(j)}(t)| \le C_j \begin{cases} t^{-1}, & 1 < t, \quad j = 1, 2, \dots, \\ \sqrt{t}, & 0 < t < 1, \quad j = 0, 1, \dots. \end{cases}$.

Then, $L_{\pm,\sigma}$ is C^∞ with respect to $\rho > 0$ and $r > 0$ and, for a constant $C > 0$,

$$|L_{\pm,\sigma}(\rho, r)| \le C\langle\rho - r\rangle^{-(\sigma+1)} \tag{6.69}$$

Proof. We prove the lemma for $L_{+,\sigma}$. The proof for $L_{-,\sigma}$ is similar. It is obvious that $L_{+,\sigma}(\rho, r)$ is smooth and is bounded for $\rho, r > 0$ and, it suffices to prove (6.69) for $|\rho - r| \ge 1$. We apply integration by parts $\sigma + 1$ times to

$$L_{+,\sigma}(\rho, r) = \frac{(-i)^{\sigma+1}}{(\rho - r)^{\sigma+1}}\int_0^\infty \left(\partial_\lambda^{\sigma+1}e^{i\lambda(\rho-r)}\right)\lambda^\sigma h_+(\lambda\rho)g_-(r\lambda)F(\lambda)d\lambda.$$

By Leibniz' rule, derivatives $(\lambda^\sigma h_+(\lambda\rho)g_-(r\lambda)F(\lambda))^{(\kappa)}$ are linear combinations over indices (β, γ, δ) such that $\kappa - \sigma \le \beta + \gamma + \delta \le \kappa$ of

$$\lambda^{\sigma-\kappa+\delta}(\lambda\rho)^\beta h^{(\beta)}(\lambda\rho)(r\lambda)^\gamma g_-^{(\gamma)}(r\lambda)F^{(\delta)}(\lambda) \tag{6.70}$$

and they converge to 0 as $\lambda \to 0$ if $\kappa \le \sigma$. It follows that $(\rho-r)^{\sigma+1}L_{+,\sigma}(\rho,r)$ is the linear combination over the same set of (β,γ,δ) as above but with $\kappa=\sigma+1$ of

$$I_{\beta\gamma\delta}(\rho,r)=\int_0^\infty e^{i(\rho-r)\lambda}\lambda^{\delta-1}(\lambda\rho)^\beta h^{(\beta)}(\lambda\rho)(r\lambda)^\gamma g_-^{(\gamma)}(r\lambda)F^{(\delta)}(\lambda)d\lambda.$$

It suffices to show that $I_{\beta\gamma\delta}(\rho,r)$ is bounded. If $\delta\neq 0$, $F^{(\delta)}(\lambda)=0$ outside $0<c_0<\lambda<c_1<\infty$ and it is clear that $I_{\beta\gamma\delta}(\rho,r)\le_{|\,\cdot\,|}C$. Thus, we assume $\delta=0$ in what follows. We may also assume $0<r<\rho<\infty$ by symmetry. We split the interval of integration as $(0,\infty)=(0,1/\rho)\cup[1/\rho,1/r]\cup(1/r,\infty)$ and denote integrals over these intervals by I_1, I_2 and I_3 in this order so that $I_{\beta\gamma\delta}(\rho,r)=I_1+I_2+I_3$.
(1) If $0<\lambda<1/\rho$ then $0<r\lambda<\rho\lambda<1$ and $(\rho\lambda)^\beta h^{(\beta)}(\rho\lambda)\le_{|\,\cdot\,|}C\sqrt{\rho\lambda}$ and $(r\lambda)^\gamma g_-^{(\gamma)}(r\lambda)\le_{|\,\cdot\,|}C\sqrt{r\lambda}$. It follows that

$$I_1\le_{|\,\cdot\,|}C\int_0^{1/\rho}\sqrt{\rho r}d\lambda=C\sqrt{\frac{r}{\rho}}\le C \tag{6.71}$$

(2) If $1/\rho\le\lambda\le 1/r$, we have $0<r\lambda\le 1\le\rho\lambda$ and we estimate as $(\rho\lambda)^\beta h^{(\beta)}(\rho\lambda)\le_{|\,\cdot\,|}C$ and $(r\lambda)^\gamma g_-^{(\gamma)}(r\lambda)\le_{|\,\cdot\,|}C\sqrt{r\lambda}$. It follows that

$$I_2\le_{|\,\cdot\,|}C\int_{1/\rho}^{1/r}\lambda^{-\frac12}\sqrt{r}d\lambda=2C\sqrt{r}\left(\frac{1}{\sqrt{r}}-\frac{1}{\sqrt{\rho}}\right)\le 2C. \tag{6.72}$$

(3) Finally if $1<r\lambda<\rho\lambda$, then we likewise estimate

$$(\lambda\rho)^\beta h^{(\beta)}(\lambda\rho)(r\lambda)^\gamma g_-^{(\gamma)}(r\lambda)\le_{|\,\cdot\,|}C\begin{cases}(r\lambda)^{-1}, & \text{if } \beta=0,\ \gamma\neq 0\\ (\rho\lambda)^{-1}, & \text{if } \beta\neq 0,\ \gamma=0,\\ (\rho\lambda)^{-1}(r\lambda)^{-1}, & \text{if } \beta,\gamma\neq 0.\end{cases}$$

The right hand side is bounded by $Cr^{-1}\lambda^{-1}$ and

$$I_3\le_{|\,\cdot\,|}C\int_{1/r}^\infty\lambda^{-2}r^{-1}d\lambda=C.$$

This completes the proof. □

Proposition 6.13. *Let $m\ge 6$ and $\phi\in\mathcal{E}$. For $1\le p\le\frac{2(m-1)}{m+1}$, we have*

$$\|Z^{\nu\nu}u\|_p\le C_p\|u\|_p. \tag{6.73}$$

Proof. We recall (6.61). Lemma 6.12 implies $L_\pm(\rho,r)\le_{|\,\cdot\,|}C\langle\rho-r\rangle^{-(m-3)}$. It follows by Young's inequality and (6.65) that

$$\|Z^{\nu\nu}u\|_p\le C\|V\phi\|_1\left(\int_0^\infty\left(\int_{\mathbb{R}}\frac{\rho^{\frac{m-1}{p}-\frac{m+1}{2}}|r^{\frac{m+1}{2}}M(r)|}{\langle\rho-r\rangle^{m-3}}dr\right)^p d\rho\right)^{\frac1p}. \tag{6.74}$$

Define $\kappa = \frac{m-1}{p} - \frac{m+1}{2}$, then $\kappa \geq 0$ for $1 \leq p \leq \frac{2(m-1)}{m+1}$ and $m-3-\kappa \geq \frac{3}{2}$ for any $1 \leq p < \infty$ if $m \geq 6$. Thus, we may estimate

$$\rho^{\kappa}\langle \rho - r\rangle^{-(m-3)} \leq C \begin{cases} \langle \rho - r\rangle^{-\frac{3}{2}} & \text{if } |r| \leq 1 \\ \langle \rho - r\rangle^{-\frac{3}{2}}|r^{\kappa}| & \text{if } |r| \geq 1 \end{cases}$$

and Young's inequality implies

$$\|Z^{\nu\nu}u\|_p \leq C\|V\phi\|_1 \left(\int_0^1 |r^{\frac{m+1}{2}}M(r)|^p dr + \int_1^\infty |r^{\frac{m-1}{p}}M(r)|^p dr\right)^{\frac{1}{p}},$$

which is bounded by $C(\|V\phi * u\|_\infty + \|V\phi * u\|_p) \leq (\|V\phi\|_{p'} + \|V\phi\|_1)\|u\|_p$. This completes the proof of the proposition. □

6.2.2 The case $\frac{2(m-1)}{m-3} \leq p < \infty$

Lemma 6.14. *Let $m \geq 6$ and $\phi \in \mathcal{E}$. Then, $Z^{\nu\nu}(\phi)$ is bounded in $L^p(\mathbb{R}^m)$ for any $\frac{2(m-1)}{m-3} \leq p < \infty$.*

Proof. we apply integration by parts to (6.60) by using the identity that $\int_{\mathbb{R}} e^{-i(1+2a)r\lambda} r^{\nu+1}M(r)dr = i(1+2a)^{-1}\partial_\lambda \left(\int_{\mathbb{R}} e^{-i(1+2a)r\lambda} r^{\nu} M(r)dr\right)'$. We see that $\rho^{-\nu}Q^{a,b}_{\nu\nu}(\rho)$ is equal to

$$\begin{aligned}&\frac{(-1)^{\nu+1}}{2\pi\rho^{\nu}(i(1+2a))}\int_0^\infty e^{i(1+2b)\rho\lambda}(\lambda^{m-3}F(\lambda))'\left(\int_{\mathbb{R}} e^{-i(1+2a)r\lambda}r^{\nu}M(r)dr\right)d\lambda \\ &+ \frac{(-1)^{\nu+1}(1+2b)}{2\pi\rho^{\nu-1}(1+2a)}\int_0^\infty e^{i(1+2b)\rho\lambda}\lambda^{m-3}F(\lambda)\left(\int_{\mathbb{R}} e^{-i(1+2a)r\lambda}r^{\nu}M(r)dr\right)d\lambda\end{aligned}$$

The argument similar to the one at the beginning of the proof of Proposition 6.13 shows that the operator produced by the first line has the same L^p property as $Z^{(\nu-1)\nu}$ and, hence, is bounded in $L^p(\mathbb{R}^m)$ for any $1 < p < \infty$. Thus, we need consider the operator produced by the second line, which we substitute for $Q^{a;b}_{\nu\nu}(\rho)/\rho^{\nu}$ in (6.8) and the resulting function for $Z^{\nu\nu}_{a,b}(\phi)u(x)$ in (6.7). The result is given by (6.61) where $X_\nu(\rho)$ is replaced by

$$\begin{aligned}\tilde{X}_\nu(\rho) = \frac{C}{\rho^{\nu-1}}\int_0^\infty &\left\{e^{i\lambda\rho}\lambda^{m-3}\left(\int_0^\infty \frac{(1+2b)e^{2i\lambda\rho b}}{(1+b)^{\nu+\frac{1}{2}}}\frac{db}{\sqrt{b}}\right)\right. \\ &\times \left.\int_{\mathbb{R}} e^{-i\lambda r}\left(\int_0^\infty \frac{(1+2a)^{-1}e^{-2ia\lambda r}}{(1+a)^{\nu+\frac{1}{2}}}\frac{da}{\sqrt{a}}\right) r^{\nu}M(r)dr\right\} F(\lambda)d\lambda, \qquad (6.75)\end{aligned}$$

which can be simplified into the form (6.65), (6.66) with the roles of g and h being replaced and the factors $\rho^{-\left(\nu+\frac{3}{2}\right)}$ and $r^{\nu+2}|r|^{-\frac{1}{2}}$ being replaced by $\rho^{-\left(\nu-\frac{1}{2}\right)}$ and $r^{\nu}|r|^{-\frac{1}{2}}$ respectively. Then, Lemmas 6.11 and 6.12 imply

$$X_\nu(\rho) \leq_{|\cdot|} \frac{C}{\rho^{\nu-\frac{1}{2}}}\int_{\mathbb{R}}\langle \rho - r\rangle^{3-m}|r|^{\nu}|r|^{-\frac{1}{2}}M(r)dr.$$

We estimate $\|X_\nu(|y|)\|_{L^p(|y|\geq 1)}$ for $p \geq \frac{2(m-1)}{m-3}$ and $\|X_\nu(|y|)\|_{L^1(|y|<1)}$. Let $\kappa = \frac{m-1}{p} - \nu + \frac{1}{2}$. If $p \geq \frac{2(m-1)}{m-3}$, then $\kappa \leq 0$ and $m-3+\kappa \geq \frac{3}{2}$ for $m \geq 6$ and for $\rho \geq 1$

$$\rho^\kappa \langle \rho - r\rangle^{3-m}|r|^{\nu-\frac{1}{2}} \leq C\langle \rho - r\rangle^{-\frac{3}{2}}\langle r\rangle^\kappa |r|^{\nu-\frac{1}{2}} \leq C\langle \rho - r\rangle^{-\frac{3}{2}}|r|^{\frac{m-1}{p}}.$$

It follows by Young's inequality that for any $2 \leq p < \infty$,

$$\begin{aligned}\|X_\nu(|y|)\|_{L^p(|y|\geq 1)} &\leq C\left(\int_0^\infty \left|\int_{\mathbb{R}} \langle \rho - r\rangle^{-\frac{3}{2}}|r|^{\frac{m-1}{p}} M(r)dr\right|^p d\rho\right)^{\frac{1}{p}} \\ &\leq C\left(\int_0^\infty |M(r)|^p r^{m-1}dr\right)^{\frac{1}{p}} \leq C\|V\phi\|_1\|u\|_p. \qquad (6.76)\end{aligned}$$

When $\rho \leq 1$, we have $\rho^{m-1-\nu+\frac{3}{2}} \leq 1$ and $\langle \rho - r\rangle^{3-m} \leq C\langle r\rangle^{3-m}$. Hence,

$$\|X_\nu(|y|)\|_{L^1(|y|<1)} \leq C\int_{\mathbb{R}} \langle r\rangle^{3-m}|r|^{\nu-\frac{1}{2}}|M(r)|dr \leq C\|M\|_\infty \leq C\|V\phi\|_{p'}\|u\|_p.$$

We therefore obtain by using Young's inequality again after splitting the integral corresponding to (6.61) into the ones over $|y| < 1$ and $|y| \geq 1$ that

$$\|Z^{\nu\nu}u\|_p \leq C(\|V\phi\|_1^2 + \|V\phi\|_p\|V\phi\|_{p'})\|u\|_p.$$

This completes the proof. $\square$

6.3 Estimate of $\|Z_{\log}u\|_p$

In this section we study $Z_{\log}$ and prove the following lemma. The combination of the lemma with results of the previous subsections proves Theorem 1.5 for even dimensions $m \geq 6$, the formal proof of which will be omitted.

Lemma 6.15. (1) *If $m = 6$, then $Z_{\log}$ is bounded in $L^p(\mathbb{R}^m)$ for all $1 < p < m$. If $\mathcal{E} = \mathcal{E}_0$, then so is $Z_{\log}$ for all $1 < p < \infty$.*

(2) *If $m \geq 8$, then $Z_{\log}$ is bounded in $L^p(\mathbb{R}^m)$ for all $1 < p < \infty$.*

Proof. We prove the lemma for $m = 6$ only. The proof for $m \geq 8$ is similar and easier. Out of three operators on the right of (3.27) for $m = 6$, we first study

$$Z_{1,\log} = \int_0^\infty G_0(\lambda)(V\varphi \otimes V\varphi)\lambda \log\lambda (G_0(\lambda) - G_0(-\lambda))F(\lambda)d\lambda, \qquad (6.77)$$

where we have ignored the constant $\omega_{m-1}/\pi(2\pi)^m$ which is not important. Since $Z_{1,\log} = 0$, if $\mathcal{E} = \mathcal{E}_0$, it suffices to prove (1) for $1 < p < \frac{m}{m-1}$ and $\frac{m}{2} < p < m$. By using (2.6) and (2.18) as previously, we express $Z_{1,\log}$ as the sum over $0 \leq j,k \leq \nu$ of

$$Z_{1,\log}^{jk}u(x) = C_{jk}T_j^{(a)}T_k^{(b)}\left[\int_{\mathbb{R}^m} \frac{(V\phi)(x-y)Q_{jk,\log}^{a,b}(|y|)}{|y|^{m-2-k}}dy\right], \qquad (6.78)$$

where $Q^{a,b}_{jk,\log}(\rho)$ are defined by (6.3) or (6.5) (for the case $j=0$) by replacing λ^{j+k-1} or λ^k respectively by $\lambda^{j+k+1}\log\lambda$ or $\lambda^{k+2}\log\lambda$. We prove

$$\|Z^{jk}_{1,\log}u\|_p \leq C\|u\|_p \tag{6.79}$$

separately for $(j,k)\neq(\nu,\nu)$ and $(j,k)=(\nu,\nu)$ by repeating the argument in corresponding subsections.

Let $(j,k)\neq(\nu,\nu)$. We first observe that, if $j\geq 1$, Fourier inverse transforms of the derivatives upto the order $k+1$ of $\lambda^{j+k+1}(\log\lambda)F$ have the RDIM

$$\mathcal{F}^*(\lambda^{j+k+1}(\log\lambda)F)^{(l)})(\rho)\leq_{|\cdot|} C(1+\rho)^{-2}\langle\log(1+\rho)\rangle, \quad 0\leq l\leq k+1$$

and estimates corresponding to (6.11) and (6.27) are satisfied by $Q^{a,b}_{jk,\log}(\rho)$ respectively for $1\leq j\leq\nu$ and for $j=0$ (without producing the boundary term). Then, the argument in §6.1.1 goes through for $Z^{jk}_{1,\log}$ and produces estimate (6.79) for $1<p<\frac{m}{m-1}$. By the same reason the estimate corresponding (6.13) for $m/2<p<m$ is satisfied by $Q^{a,b}_{jk,\log}(\rho)$ for all j,k and we likewise have (6.79) for $m/2<p<m$ by using the argument of the first part of proof of Lemma 6.8. It is then obvious that the same holds for $Z_{2,\log}$ which is obtained from $Z_{1,\log}$ by replacing $\lambda\log\lambda$ by $\lambda^3(\log\lambda)$ and, that the operator

$$Z^{(a,b)}_{3,\log} = \int_0^\infty G_0(\lambda)(\varphi_a\otimes\psi_b)\lambda^3(\log\lambda)^2(G_0(\lambda)-G_0(-\lambda))F(\lambda)d\lambda. \tag{6.80}$$

produced by $\lambda^2\log\lambda F_2$ of (3.18) satisfies (6.79) for all $1<p<m$.

We next prove (6.79) when $(j,k)=(\nu,\nu)$. It suffices prove it for $1<p<p_0$ for some $p_0>1$ and $p\geq p_1$ for some $p_1>2$. The argument at the beginnings of §6.2.1 and §6.2.2 shows that respectively for $1<p<p_0$ and $p\geq p_1$, we have only to estimate operators obtained by replacing $Q^{a,b}_{jk,\log}(\rho)$ by

$$\frac{1+2a}{(1+2b)\rho^{\nu+1}}\int_0^\infty e^{i(1+2b)\rho\lambda}\lambda^{m-1}\log\lambda F(\lambda)\left(\int_{\mathbb{R}} e^{-i(1+2a)r\lambda}r^{\nu+2}M dr\right)d\lambda$$

and

$$\frac{1+2b}{(1+2a)\rho^{\nu-1}}\int_0^\infty e^{i(1+2b)\rho\lambda}\lambda^{m-1}\log\lambda F(\lambda)\left(\int_{\mathbb{R}} e^{-i(1+2a)r\lambda}r^{\nu+2}M dr\right)d\lambda$$

in (6.78). We then repeat the argument of §6.2. We have $\lambda^{m-2}\log\lambda$ in place of λ^{m-4} in (6.66). If we change λ^σ by $\lambda^{\sigma+2}\log\lambda$ in the definition (6.68) of $\tilde{L}_\pm(\rho,r)$, then (6.69) is satisfied with faster decaying factor $\langle\rho-r\rangle^{-(\sigma+2)}$ in place of $\langle\rho-r\rangle^{-(\sigma+1)}$. Thus, $\|Z^{\nu\nu}_{\log}u\|_p$ is bounded $C\|V\phi\|_1$ times (6.74) with $\langle\rho-r\rangle^{-(m-2)}$ in place of $\langle\rho-r\rangle^{-(m-3)}$ and this proves the lemma for $1<p<p_0$. The proof for $p\geq p_1$ is similar and we omit further details. $\square$

REFERENCES

[1] S. Agmon, *Spectral properties of Schrödinger operators and scattering theory*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 2 (1975), 151–218.

[2] M. Aizenman and B. Simon, *Brownian motion and Harnack inequality for Schrödinger equations*, Comm. Pure Appl. Math. 35 (1982).

[3] G. Artbazar and K. Yajima, *The L^p-continuity of wave operators for one dimensional Schrödinger operators*, J. Math. Sci. Univ. Tokyo 7 (2000), 221-240.

[4] M. Beceanu, *Structure of wave operators for a scaling-critical class of potentials*, Amer. J. Math. 136 (2014), 255–308.

[5] J. Bergh and J. Löfström, *Interpolation spaces, an introduction*, Springer Verlag, Berlin-Heidelberg-New York (1976).

[6] P. D'Ancona and L. Fanelli, *L^p–boundedness of the wave operator for the one dimensional Schrödinger operator*, Commun. Math. Phys. 268 (2006), 415–438.

[7] M. B. Erdoğan and W. Schlag, *Dispersive Estimates for Schrödinger Operators in the Presence of a Resonance and/or an Eigenvalue at Zero Energy in Dimension Three: I*, Dyn. Partial Differ. Equ. 1 (2004), 359–379.

[8] D. Finco and K. Yajima, *The L^p boundedness of wave operators for Schrödinger operators with threshold singularities. II. Even dimensional case.* J. Math. Sci. Univ. Tokyo 13 (2006), 277–346.

[9] A. Galtbayer and K. Yajima *Resolvent estimates in amalgam spaces and asymptotic expansions for Schrödinger equations*, J. Math. Soc. Japan

[10] M. Goldberg and W. Green, *The L^p boundedness of wave operators for Schrödinger operators with threshold singularities*, arXiv:1508.06300.

[11] L. Grafakos, *Modern Fourier analysis*, Springer Verlag, New York (2009).

[12] R. L. Graham, D. E.Knuth and O. Patashnik, *Concrete matematics*, Addison-Wesley Pub. Co. New York (1988).

[13] A. Jensen, *Spectral properties of Schrödinger operators and time decay of the wave functions, Results in $L_2(\mathbb{R}^m)$, $m \geq 5$*, Duke Math. J. 47 (1980), 57–80.

[14] A. Jensen and T. Kato, *Spectral properties of Schrödinger operators and time-decay of the wave functions*, Duke Math. J. 46 (1979), 583–611.

[15] A. Jensen and K. Yajima *A remark on L^2-boundedness of wave operators for two dimensional Schrödinger operators*, Commun. Math. PHys. 225 (2002), 633–637.

[16] A. Jensen and K. Yajima, *On L^p-boundedness of wave operators for 4-dimensional Schrödinger with threshold singularities*, Proc. Lond. Math. Soc. (3) 96 (2008), 136–162.

[17] T. Kato, *Growth properties of solutions of the reduced wave equation with a variable coefficient*, Comm. Pure Appl. Math 12 (1959), 403–425.

[18] T. Kato, *Wave operators and similarity for non-selfadjoint operators*, Ann. Math. 162 (1966), 258–279.

[19] S. T. Kuroda, *Introduction to Scattering Theory*, Lecture Notes, Aarhus University

[20] M. Murata, *Asymptotic expansions in time for solutions of Schrödinger-type equations*, J. Funct. Anal., 49 (1982), 10–56.

[21] M. Reed and B. Simon, *Methods of modern mathematical physics vol II, Fourier analysis, selfadjointness*, Academic Press, New-York, San Francisco, London (1975).

[22] M. Reed and B. Simon, *Methods of modern mathematical physics vol III, Scattering theory*, Academic Press, New-York, San Francisco, London (1979).

[23] M. Reed and B. Simon, *Methods of modern mathematical physics vol IV, Analysis of Operators*, Academic Press, New-York, San Francisco, London (1978).

[24] K. Stempakit Transplantation theorems, survey, J. Fourier Anal. Appl. 17 (2011), 408–430.

[25] E. M. Stein, *Singular integrals and differentiability properties of functions*, Princeton Univ. Press, Princeton, NJ. (1970)

[26] E. M. Stein, *Harmonic analysis, real-variable methods, orthogonality, and oscillatory integrals*, Princeton Univ. Press, Princeton, NJ. (1993).

[27] R. Weder, *L^p-$L^{p'}$ estimates for the Schrödinger equations on the line and inverse scattering for the nonlinear Schrödinger equation with a potential*, J. Funct. Anal. 170 (2000), 37–68.

[28] E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis*. Cambridge University Press; 4th edition (1927).

[29] K. Yajima, *The $W^{k,p}$-continuity of wave operators for Schrödinger operators*, J. Math. Soc. Japan 47 (1995), 551-581.

[30] K. Yajima, *The $W^{k,p}$-continuity of wave operators for Schrödinger operators III*, J. Math. Sci. Univ. Tokyo 2 (1995), 311–346.

[31] K. Yajima, *The L^p-boundedness of wave operators for two dimensional Schrödinger operators*, Commun. Math. Phys. 208 (1999), 125–152.

[32] K. Yajima, *Dispersive estimate for Schrödinger equations with threshold singularities*, Comm. Math. Phys. 259 (2005), no. 2, 475–509.

[33] K. Yajima, *The L^p boundedness of wave operators for Schrödinger operators with threshold singularities I, Odd dimensional case*, J. Math. Sci. Univ. Tokyo 7 (2006), 43–93.

[34] K. Yajima, *Wave operators for Schrödinger operators with threshold singularities,revisited*, arXiv:1508.05738.

K. Yajima
Department of Mathematics
Gakushuin University
1-5-1 Mejiro
Toshima-ku
Tokyo 171-8588
Japan

Hyperelliptic Schottky Problem and Stable Modular Forms

Giulio Codogni

Received: August 26, 2014
Revised: February 15, 2016

Communicated by Gavril Farkas

Abstract. It is well known that, fixed an even, unimodular, positive definite quadratic form, one can construct a modular form in each genus; this form is called the theta series associated to the quadratic form. Varying the quadratic form, one obtains the ring of stable modular forms. We show that the differences of theta series associated to specific pairs of quadratic forms vanish on the locus of hyperelliptic Jacobians in each genus. In our examples, the quadratic forms have rank 24, 32 and 48. The proof relies on a geometric result about the boundary of the Satake compactification of the hyperelliptic locus. We also study the monoid formed by the moduli space of all principally polarised abelian varieties, the operation being the product of abelian varieties. We use this construction to show that the ideal of stable modular forms vanishing on the hyperelliptic locus in each genus is generated by differences of theta series.

2010 Mathematics Subject Classification: 14H42 , 32G20
Keywords and Phrases: Schottky problem, Satake compactification, theta series

1. Introduction

The hyperelliptic Schottky problem is to characterise the locus of Jacobians of hyperelliptic curves inside the moduli space of principally polarised abelian varieties. A classical approach is to look for modular forms vanishing along the hyperelliptic locus; in other words, one looks for the equations of the hyperelliptic locus inside the moduli space of principally polarised abelian varieties.
A special kind of modular forms are the theta constants; these were used by Mumford to give a solution to the hyperelliptic Schottky problem, as reviewed in Theorem 4.4.

In this paper, we deal with stable modular forms. One nice feature of these forms is that they relate the theory of moduli spaces with the theory of quadratic forms.

To start with, let us explain what we mean by stable modular forms. Let $\mathcal{A}_g$ be the moduli space of principally polarised abelian g-fold defined over the field of complex numbers. We consider the Satake compactification $\mathcal{A}_g^S$ of $\mathcal{A}_g$. This comes with a stratification

$$\mathcal{A}_g^S = \mathcal{A}_g \sqcup \mathcal{A}_{g-1} \cdots \mathcal{A}_1 \sqcup \mathcal{A}_0$$

In particular, we have a closed embedding

$$\iota_g \colon \mathcal{A}_{g-1}^S \hookrightarrow \mathcal{A}_g^S$$

The collection of the moduli spaces $\mathcal{A}_g^S$ and these maps form a direct system of varieties; we can thus consider the ind-scheme

$$\mathcal{A}_\infty := \lim_g A_g^S$$

The basic definitions about ind-schemes are recalled in Section 2

Stable modular forms are naturally defined on $\mathcal{A}_\infty$. A stable modular form F is a collection of modular forms $(F_g)_{g\geq 0}$: each F_g is modular form on $\mathcal{A}_g^S$ and $\iota_g^* F_g = F_{g-1}$. We will recall the theory in Section 3, in particular see Definition 3.3. A classical and surprising fact is that we can construct a stable modular form out of an even, unimodular, positive definite quadratic form Q. This stable modular form is called theta series associated to Q, and is denoted by Θ_Q; see Definition 3.6. In particular, for any g, $\Theta_{Q,g}$ is a modular form on $\mathcal{A}_g^S$. In [Fre77], Freitag showed that all stable modular forms are linear combinations of theta series.

In this set up, we can consider the ideal of stable modular forms vanishing on the locus Hyp_g of hyperelliptic Jacobians in every genus. Let us formalise this with definition.

Definition 1.1 (Stable Equation). *A stable equation for the hyperelliptic locus is a stable modular form $(F_g)_{g\geq 0}$ such that F_g vanishes along the hyperelliptic locus Hyp_g for every g.*

Our first result is the following:

Theorem 1.2 (= Theorem 4.2). *The ideal of stable equations of the hyperelliptic locus is generated by differences of theta series*

$$\Theta_P - \Theta_Q$$

where P and Q are even, unimodular, positive definite quadratic forms of the same rank.

A key ingredient in the proof of this results is a natural monoidal structure that one can put on $\mathcal{A}_\infty$. Given two principally polarised abelian varieties, their product is still a principally polarised abelian variety but of higher dimension. This defines an operation

$$m \colon \mathcal{A}_\infty \times \mathcal{A}_\infty \to \mathcal{A}_\infty$$

The pull-back m^* gives to the ring of stable modular forms the structure of a co-commutative co-algebra. Because of this, we can run the general machinery explained in Section 2 to prove Theorem 1.2.

So far, the ideal of stable equations for the hyperelliptic locus could be trivial. Indeed, this is the case for the moduli space of curves: in [CSB14], it is shown that the ideal of stable modular form vanishing on the Jacobian locus in any genus is trivial. In other words, given a non-zero stable modular form F, there exists a g such that F_g does not vanish on the moduli space of genus g curves. In [SB13], it is similarly shown that the ideal of stable equation for the n-gonal locus, with $n \geq 3$, is trivial. However, as we are going to see, the ideal of stable equations for the locus of Jacobians of hyperelliptic curves is far from being trivial.

The first stable equation for the hyperelliptic locus was discovered by C. Poor ([Poo96]): it is the difference of the theta series associated to the quadratic forms D_{16}^+ and $E_8 \oplus E_8$; this modular form is also called the called Schottky form. To construct new stable equations, we need to know more about the geometry of the Satake compactification. The Satake compactification Hyp_g^S of Hyp_g will be defined in Section 4. We denote by $\mathcal{A}_g^{ind}$ the moduli space of indecomposable principally polarised abelian g-fold.

Theorem 1.3 (= Theorem 4.3; Transversality). *The intersection of the Stake compactification Hyp_{g+1}^S and $\mathcal{A}_g^{ind}$ inside $\mathcal{A}_{g+1}^S$ is scheme theoretically equal to Hyp_g.*

The statement was well-known at the level of sets; we call it a transversality result because it states that the scheme structure of the intersection is the reduced one. The analogue result does not hold for the moduli of curves [CSB14, Theorem 1.1] and for the moduli of n-gonal curves [SB13], with $n \geq 3$. In those cases, the failure of the transversality implies that there are no stable equations; in the hyperelliptic case, this transversality result is key in the construction of stable equations.
Combining Theorem 1.3 and Criterion 5.2 we can prove the following

Theorem 1.4 (= Corollary 5.7 and Theorem 6.1). *The difference of theta series*

$$\Theta_P - \Theta_Q$$

is a stable equation for the hyperelliptic locus when one of the following hold:

(1) $\mathrm{rk}(P) = \mathrm{rk}(Q) = 24$ *and the two quadratic forms have same number of vectors of norm* 2*;*
(2) $\mathrm{rk}(P) = \mathrm{rk}(Q) = 32$ *and the two quadratic forms do not have any vector of norm* 2*;*
(3) $\mathrm{rk}(P) = \mathrm{rk}(Q) = 48$ *and the two quadratic forms do not have any vector of norm* 2 *or* 4*;*

Each item of Corollary 1.4 concerns a finite positive number of pairs of quadratic forms. In [Kin03], it is shown that there are more than ten millions of quadratic forms meeting the hypothesis of the second item. In the first item, the "slope" of the quadratic form, i.e. the ratio between the rank and the norm of the shortest non-zero vector, is strictly bigger than the slope of the hyperelliptic locus; because of this, in the proof we need to use some non-trivial arithmetic properties of the quadratic forms: namely we use Theorem 6.2 via Corollary 6.3.

We think that the ideal of stable equations defines scheme theoretically the hyperelliptic locus inside the moduli space of indecomposable principally polarised abelian varieties in any genus. This, in particular, would imply that there are infinitely many pairs of quadratic forms which give stable equations for the hyperelliptic locus. In order to give a characterisation of these pairs, we think one should relate theta series to partition functions, as partially suggested in [GV09], [GKV10] and [Mat15].

Acknowledgement. This work is part of my Ph.D. thesis [Cod14], I would like to thank my advisor Nick Shepherd-Barron, without whom this work would not have been possible. It is a pleasure to thank Riccardo Salvati Manni for many useful advices. I also had the benefit of conversations with Enrico Arbarello, Gavril Farkas, Sam Grushevsky, Marco Matone, Sara Perna, Filippo Viviani and Roberto Volpato on the topics of this paper. I would like to thank the refree for useful comments.
My Ph.D. was funded by the DPMMS, the EPSRC and Selwyn College. I have been also supported by the Physics department of the University of Padova during two visits. Now, I am funded by the grant FIRB 2012 "Moduli Spaces and their applications" and by the ERC StG 307119 - StabAGDG.

2. Ind-varieties and commutative Monoids

In this section, we recall some general definitions and results about ind-varieties and monoid. A reference about ind-variety is [Kum02, Chapter IV]. An ind-variety X is a collection $(X_n)_{n\geq 0}$ of algebraic varieties and a collection of closed embeddings

$$\iota_n \colon X_{n-1} \hookrightarrow X_n$$

We write

$$X = \lim_n X_n$$

This limit exists in the category of locally ringed spaces; however, we prefer to enlarge the category of schemes including all direct systems. This means that for us an ind-variety is a direct system of algebraic varieties.
A line bundle L on X is the data of a line bundle L_n on each X_n such that $\iota_n^* L_n = L_{n-1}$. A section s of L is a collection of sections $(s_n)_{n\geq 0}$ such that s_n is a section of L_n on X_n and the restriction of s_n to X_{n-1} is s_{n-1}. We assume that the vector space $H^0(X, L^k)$ is finite dimensional for every k. The ring of sections of L is thus defined as a projective limit in the category of graded rings

$$\mathcal{R}(X, L) := \varprojlim \mathcal{R}(X_n, L_n)\,.$$

We do not have to worry about the topology of this ring because of graded pieces $H^0(X, L^k)$ are finite dimensional. In other words, elements on $\mathcal{R}(X, L)$ are not formal power series.

Remark 2.1 (Ampleness on ind-varieties). The concept of ampleness for a line bundle L on an ind-variety X is subtle and, to the best of our knowledge, there is not a standard definition. A first definition could be that there exists a k such that L_n^k is very ample on X_n for every n. Remark that k does not depend on n. If $H^0(X, L^k)$ is finite dimensional for every k but the dimension of X_n tends to infinity when n grows, L can not be ample. A weaker definition is to ask that for every n there exists a $k = k(n)$ such that L_n^k is very ample on X_n. The example that we will study in this paper is ample just in the sense of the second definition. This second definition does not imply the classical consequences of ampleness: for instance, in this generality, it is not even clear that an ample line bundle is effective.

An ind-monoid is an ind-variety M with an associative multiplication and an identity element 1_M. A multiplication m is a family of maps

$$m_{g,h} \colon M_g \times M_h \to M_{g+h}$$

compatible with the restrictions. M is commutative if the multiplication is.

Definition 2.2 (Split monoid). *Let M be a commutative ind-monoid and L a line bundle on M. We say that M is split with respect to L if the following two conditions hold:*

(1) For every g and h

$$m_{g,h}^* L_{g+h} \cong pr_1^* L_g \otimes pr_2^* L_h =: L_g \boxtimes L_h$$

where pr_i are the projections;

(2) for every k, the vector space $H^0(M, L^k)$ is finite dimensional and spanned by characters, where a section χ of L is a character if

$$m_{g,h}^* \chi_{g+h} = \chi_g \boxtimes \chi_h \qquad \forall\, g, h\,.$$

In the language of Hopf algebras, condition (1) means that the pull-back m^* is a co-commutative co-multiplication for $\mathcal{R}(M, L)$. The definition of character makes sense only if condition (1) holds. With a slight abuse of notations, we will speak about characters of M rather than characters of the co-algebra $\mathcal{R}(M, L)$, and we will write $\chi(\alpha\beta) = \chi(\alpha)\chi(\beta)$ instead of $m_{g,h}^* \chi_{g+h}(\alpha \times \beta) = \chi_g(\alpha) \boxtimes \chi_h(\beta)$.

Lemma 2.3. Let M be a commutative monoid, suppose it is split with respect to a line bundle L, then, any set of characters is linearly independent.

Proof. This proof is standard. We argue by contradiction. Take n minimal such that there exist n linearly dependent characters $\chi_1, \dots, \chi_n$. We can write

$$\chi_n = \sum_{i=1}^{n-1} \lambda_i \chi_i \qquad \lambda_i \in \mathbb{C}\,.$$

Pick $\alpha \in M$ such that $\chi_1(\alpha) \neq \chi_n(\alpha)$. For any $\beta \in M$ we have

$$\sum_{i=1}^{n-1} \lambda_i \chi_i(\alpha)\chi_i(\beta) = \chi_n(\alpha)\chi_n(\beta) = \chi_n(\alpha)\left(\sum_{i=1}^{n-1} \lambda_i \chi_i(\beta)\right)$$

Since β is arbitrary we get

$$\sum_{i=i}^{n-1} \lambda_i(\chi_i(\alpha) - \chi_n(\alpha))\chi_i = 0\,.$$

The coefficient $\chi_1(\alpha) - \chi_n(\alpha)$ is non-zero, so we have written a non-trivial linear relation among fewer than n characters. This contradicts the minimality of n. $\square$

Proposition 2.4. *Let M be a commutative ind-monoid and N a submonoid. Suppose that M is split with respect to a line bundle L. Then the ideal I_N in $\mathcal{R}(M, L)$ of sections vanishing on N is generated by differences of characters*

$$\chi_i - \chi_j$$

Proof. Take s in I_N. We can assume that s is homogeneous and write it as a linear combination

$$s = \lambda_1\chi_1 + \cdots + \lambda_n\chi_n$$

where χ_i are characters and λ_i are constants. Restricting χ_i to N some of them might become equal. Up to relabelling the χ_i, we can fix integers $0 = m_0 < m_1 < \cdots < m_k = n$ and distinct characters $\theta_1, \ldots, \theta_k$ of N such that

$$\chi_i \mid_N = \theta_j \quad \iff \quad m_{j-1} < i \leq m_j$$

For $j = 1, \ldots, k$, let us define

$$\mu_j := \sum_{i=m_{j-1}+1}^{m_j} \lambda_i\,.$$

By hypothesis we know that

$$0 = s \mid_N = \sum_{j=1}^{k} \mu_j\theta_j\,.$$

By Lemma 2.3 we have $\mu_j = 0$ for every j, so

$$s = s - \sum_{j=1}^{k} \mu_j\chi_{m_j} = \sum_{j=1}^{k} \sum_{i=m_{j-1}+1}^{m_j} \lambda_i(\chi_i - \chi_{m_j})$$

The differences $\chi_i - \chi_{m_j}$ vanish on N for $m_{j-1} < i \leq m_j$, so we have just expressed s as linear combination of differences of characters vanishing on N. $\square$

The previous argument actually shows that every element of the ideal can be written as a linear combination of differences of characters. These results are special cases of a more general theory of Milnor and Moore. They have many applications in the study of moduli spaces, e.g. [GHT14].

3. Satake compactification, modular forms and theta series

We recall some facts about modular forms and the Satake compactification of $\mathcal{A}_g$. General references about modular forms are [BvdGHZ08] and [Mum07]. The Satake compactification was first defined in [Sat56]; a comprehensive reference is [Fre83].
The line bundle L_g of weight one modular forms on $\mathcal{A}_g$ is defined as the determinant of the Hodge bundle; it is ample and it generates the rational Picard group.

Definition 3.1 (Siegel modular form). *A weight k and degree g Siegel modular form is a section of L_g^k on $\mathcal{A}_g$.*

The universal cover of $\mathcal{A}_g$ is the Siegel upper half space $\mathcal{H}_g$; the symplectic group $Sp(2g,\mathbb{Z})$ acts on $\mathcal{H}_g$ and

$$\mathcal{A}_g = \mathcal{H}_g/Sp(2g,\mathbb{Z})$$

The line bundle L_g is trivial when it s pulled back to $\mathcal{H}_g$; therefore a modular form can be also defined as a holomorphic function on $\mathcal{H}_g$ which transforms appropriately under the action of $Sp(2g,\mathbb{Z})$.
The Satake compactification $\mathcal{A}_g^S$ is a normal projective variety defined as follows

$$\mathcal{A}_g^S := \mathrm{Proj}(\bigoplus_{n\geq 0} H^0(\mathcal{A}_g, L_g^n))$$

This is the compactification "seen" by modular forms. The line bundle L_g extends naturally to $\mathcal{A}_g^S$ because it is the $\mathcal{O}(1)$ of this Proj. For the same reason, all modular forms extend to $\mathcal{A}_g^S$

Definition 3.2 (The Siegel operator). *The Siegel operator Φ is a map of graded rings*

$$\Phi\colon \bigoplus_{n\geq 0} H^0(\mathcal{A}_g, L_g^n) \to \bigoplus_{n\geq 0} H^0(\mathcal{A}_{g-1}, L_{g-1}^n)$$

defined as

$$\Phi(F)(\tau) := \lim_{t\to+\infty} F(\tau \oplus it)\,,$$

where τ is an element of $\mathcal{H}_{g-1}$ and $t \in \mathbb{R}$. Here, we are thinking at F as a holomorphic function on $\mathcal{H}_g$.

Clearly, there is some work to do to show that $\Phi(F)$ is a well defined element of $H^0(\mathcal{A}_{g-1}, L_g^n)$; the interested reader cal look at [Fre83].
The Siegel operator is surjective for n even and larger than $2g$ ([Fre83] page 64); this means that the Siegel operator defines a closed embedding of $\iota_g\colon \mathcal{A}_{g-1}^S \hookrightarrow \mathcal{A}_g^S$. One can check that the image of $\mathcal{A}_{g-1}^S$ is the boundary $\partial\mathcal{A}_g^S$ of $\mathcal{A}_g^S$, so we obtain a stratification

$$\mathcal{A}_g^S = \mathcal{A}_g \sqcup \mathcal{A}_{g-1}^S = \mathcal{A}_g \sqcup \mathcal{A}_{g-1} \cdots \mathcal{A}_1 \sqcup \mathcal{A}_0\,.$$

By construction, the pull-back $\iota_g^* L_g$ is isomorphic to L_{g-1}, and the pull-back $\iota_g^*\colon H^0(\mathcal{A}_g, L_g^n) \to H^0(\mathcal{A}_{g-1}, L_{g-1}^n)$ is the Siegel operator. Again, a reference is [Fre83].

The system of varieties $\mathcal{A}_g^S$ together with the closed embeddings ι_g induced by the Siegel operators forms a direct system, so we can define the ind-variety

$$\mathcal{A}_\infty := \lim_g \mathcal{A}_g^S$$

We follow the notations of Section 2. The line bundles L_g define a line bundle

$$L_\infty := \lim_g L_g$$

on $\mathcal{A}_\infty$. This line bundle is called the line bundle of weight one stable stable modular forms.

Definition 3.3 (Stable modular forms). *A weight k stable modular form F is a section of L_∞^k. More concretely, it is a collection*

$$F = (F_g)_{g\geq 0}$$

where F_g is a modular form of weight k on $\mathcal{A}_g$ and

$$\Phi(F_{g+1}) = F_g$$

Recall that each line bundle L_g is ample on $\mathcal{A}_g$; however, the same assertion is problematic for L_∞, as explained in Remark 2.1.
We now define a structure of commutative monoid on $\mathcal{A}_\infty$. Given two principally polarised abelian varieties of dimension respectively g and h, their product is still a principally polarised abelian variety of dimension $g+h$. This gives a commutative operation

$$\begin{array}{rccc} m\colon & \mathcal{A}_\infty \times \mathcal{A}_\infty & \to & \mathcal{A}_\infty \\ & ([X],[Y]) & \mapsto & [X\times Y] \end{array}$$

The identity element is $\mathcal{A}_0$.

Lemma 3.4. On $\mathcal{A}_\infty \times \mathcal{A}_\infty$, we have

$$m^* L_\infty = L_\infty \boxtimes L_\infty\,.$$

Proof. For every pair of integers g and h one looks at the morphism

$$m\colon \mathcal{A}_g \times \mathcal{A}_h \to \mathcal{A}_{g+h}$$

The fibre of the Hodge bundle E_g at a point $[X]$ of $\mathcal{A}_g$ is the tangent space at the identity of X. This implies that m^*E_{g+h} is $E_g \boxtimes E_h$. The statement on L_g follows by taking the determinant. □

We have the following useful formal consequence

Proposition 3.5. *The pull-back m^* defines a co-commutative co-multiplication on the algebra of stable modular forms $\mathcal{R}(\mathcal{A}_\infty, L_\infty)$.*

The ring of stable modular form, so far, could be trivial. However, there is a classical and surprising way to produce plenty of stable modular forms out of quadratic forms. Let us go trough all definitions. A quadratic form is a pair (Λ, Q), where Λ is a finitely generated free group and Q is a $\mathbb{Z}$-valued bilinear form on Λ. The rank of the quadratic form is defined as the rank of Λ. The elements of Λ are called vectors, and the norm of a vector v is $Q(v,v)$. We always assume Q to be even (i.e. $Q(v,v)$ is even for every v), unimodular (i.e. $\det Q = 1$) and positive definite. Often, we will denote a quadratic form just by Q.

Definition 3.6 (Theta series). *Let (Λ, Q) be an even unimodular positive definite quadratic form and g a positive integer, the associated theta series is*

$$\Theta_{Q,g}(\tau) := \sum_{x_1,\ldots,x_g\in\Lambda} \exp(\pi i \sum_{i,j} Q(x_i,x_j)\tau_{ij})$$

where τ belongs to $\mathcal{H}_g$.

This is a weight $\frac{1}{2}\operatorname{rk}(\Lambda)$ and degree g modular form. By explicit computation, one sees that the Siegel operator 3.2 acts as follows

$$\Phi(\Theta_{Q,g+1}) = \Theta_{Q,g}\,,$$

so the collection of all theta series

$$\Theta_Q := (\Theta_{Q,g})_{g\geq 0}$$

is a stable modular form. Given $X \in \mathcal{A}_g$ and $Y \in \mathcal{A}_h$, we have the factorisation property

$$\Theta_{Q,g+h}([X\times Y]) = \Theta_{Q,g}(X)\Theta_{Q,h}(Y)\,,$$

which means that the *theta series are characters* for the monoid $\mathcal{A}_\infty$.

Example 3.7 (Quadratic forms and theta constants). In some cases, theta series can be written in term of theta constants, let us give some examples following [Igu81]. Let E_8 be the quadratic form associate to the Dynkin diagram E_8. Using a similar definition, for every integer k one can define the Witt quadratic forms W_{8k}. The quadratic form W_{8k} has rank $8k$, it is equal to E_8 for $k=1$ and to D_{16}^+ for $k=2$. Up to a constant, we have the following expansion

$$\Theta_{W_{8k},g}(\tau) = \sum_{\epsilon \text{ even}} \theta[\epsilon]^{8k}(\tau)\,,$$

where the sum runs over all the even theta characteristics. In particular, the well-known Schottky form can be written as

$$\Theta_{D_{16}^+} - \Theta_{E_8\oplus E_8} = 2^{-g}\sum_{\epsilon \text{ even}} \theta[\epsilon]^{16}(\tau) - 2^{-2g}(\sum_{\epsilon \text{ even}} \theta[\epsilon]^{8}(\tau))^2$$

In general, a theta series will not have such a simple expression in term of theta constants. In [SM89, Section 3], there is a systematic analysis of the theta series which can be expanded in this way; the results of that paper relies upon [Mum07, Theorem 6.3].

The ring of stable modular forms is described by the following result of Freitag:

Theorem 3.8 (Theorem 2.5 of [Fre77]). *The ring of stable modular forms $\mathcal{R}(\mathcal{A}_\infty, L_\infty)$ is the polynomial ring in the theta series associated to irreducible quadratic forms.*

Freitag's main contribution was to show that $H^0(\mathcal{A}_\infty, L_\infty^k)$ is spanned by theta series for every k. There are finitely many quadratic forms of a given rank, so we already learn that $H^0(\mathcal{A}_\infty, L_\infty^k)$ is finite dimensional. This result, together with Lemma 3.4 and the fact that theta series are characters, means that $\mathcal{A}_\infty$ equipped with the line bundle L_∞ is a split monoid in the sense of Definition 2.2. Freitag's claim about the polynomial structure now follows easily from Proposition 2.4.

4. Satake compactification of the hyperelliptic locus

In this section we define the Satake compactification of the hyperelliptic locus and we prove Theorems 1.2 and 1.3. Consider the Jacobian morphism

$$j\colon Hyp_g \to \mathcal{A}_g$$

mapping a curve to its Jacobian.

Definition 4.1 (Satake compactification). *The Satake compactification Hyp_g^S of the hyperelliptic locus Hyp_g is the scheme-theoretic closure of $j(Hyp_g)$ inside $\mathcal{A}_g^S$.*

A degeneration of a hyperelliptic Jacobian is still the Jacobian of a hyperelliptic curve ([Hoy63]), so we have a stratification

$$Hyp_g^S = \bigsqcup_{\sum g_i \leq g} Hyp_{g_1} \times \cdots \times Hyp_{g_k}$$

Equivalently, the Satake compactification is the image of the Deligne-Mumford compactification of Hyp_g under the morphism which maps a curve to the Jacobian of its normalisation.

In particular, Hyp_{g+1}^S contains Hyp_g^S as a scheme, so we can define the commutative ind-monoid

$$Hyp_\infty := \lim_g Hyp_g^S$$

Using the monoid structure we can show the following

Theorem 4.2. *The ideal of stable modular forms vanishing on Hyp_∞ is generated by differences of theta series*

$$\Theta_P - \Theta_Q$$

where P and Q are even, unimodular, positive definite quadratic forms of the same rank.

Proof. We know that Hyp_∞ is a commutative sub-monoid of $\mathcal{A}_\infty$ and $\mathcal{A}_\infty$ satisfies the hypotheses of Definition 2.2. The theta series are the characters of co-algebra $\mathcal{R}(\mathcal{A}_\infty, L_\infty)$, so the result is a direct consequence of Proposition 2.4. □

So far, the ideal studied in the theorem could be trivial. A key tool to show that a modular form is a stable equation for the hyperelliptic locus is the following geometric result. Let $\mathcal{A}_g^{ind}$ be the moduli space of indecomposable principally polarised abelian g-fold.

Theorem 4.3 (Transversality). *Inside $\mathcal{A}_{g+1}^S$, the intersection of Hyp_{g+1}^S and $\mathcal{A}_g^{ind}$ is scheme theoretically equal to Hyp_g.*

The statement was well-known at the level of sets (cf [Hoy63] or [ACG11, Lemma 11.6.14]); we call it a transversality result because it states that the scheme structure of the intersection is the reduced one.

Proof. We work with level structure $(4,8)$, and we denote by $\mathcal{A}_g^S(4,8)$ the Satake compactification of the moduli space $\mathcal{A}_g(4,8)$ of principally polarised abelian g-fold with level structure $(4,8)$. This amounts to take a finite Galois cover of $\mathcal{A}_g^S$. Now, $\mathcal{A}_{g+1}^S(4,8)$ has several boundary components, all isomorphic to $\mathcal{A}_g^S(4,8)$. We will fix one of them, let us call it V. The hyperelliptic locus $Hyp_{g+1}^S(4,8)$ breaks into several irreducible components (cf. [Tsu91]); an irreducible component is identified by the choice of a fundamental system $\mathfrak{m}$ of theta characteristic, we will fix such an $\mathfrak{m}$ and denote by $Y = Y_{\mathfrak{m}}$ the corresponding irreducible component. Since the cover is Galois, locally the intersection of Hyp_{g+1}^S and $\mathcal{A}_g$ is isomorphic to the intersection of Y and V. Because of this, it is enough to show that the scheme-theoretic intersection of V^{ind} and Y is reduced. We need to work at level $(4,8)$ to apply the following result, which is due set-theoretically to Mumford and scheme theoretically to Salvati Manni.

Theorem 4.4 ([SM03]). *Fix a fundamental system of theta characteristic $\mathfrak{m} = (m_1, \ldots, m_{2g+1})$, and let b be the sum of the odd m_i. Then, the corresponding irreducible component $Y = Y_{\mathfrak{m}}$ of $Hyp_g(4,8)$ is scheme theoretically defined by the vanishing of the theta constants θ_{m+b} such that $m = m_{i_1} + \cdots + m_{i_k}$ for $k < g$ and $k \equiv g, g+1$, and the non-vanishing of the remaining theta constants.*

In the statement, the non-vanishing of the remaining theta constants is needed to rule out the loci of decomposable abelian varieties; our statement is about Jacobians of smooth curves, so we are already working outside these loci. Let $\mathfrak{m}$ be a $g+1$ dimensional system of fundamental theta characteristic, e.g. the one defined in equation 7 of [SM03]. Let $Y = Y_{\mathfrak{m}}$ be the corresponding irreducible component of $Hyp_{g+1}^S(4,8)$. Let $I_{g+1}(Y_{\mathfrak{m}})$ be the ideal of modular forms generated by the theta constants vanishing along $Y_{\mathfrak{m}}$. Because of Salvati Manni's result, this ideal defines scheme-theoretically $Y_{\mathfrak{m}}$. By direct computation one sees that

$$\Phi(\theta \begin{bmatrix} \epsilon & 0 \\ \epsilon' & \delta \end{bmatrix}) = \theta \begin{bmatrix} \epsilon \\ \epsilon' \end{bmatrix}$$

for δ equal either to 0 or 1. In the previous formula, Φ is the Siegel operator, i.e. the restriction operator from $\mathcal{A}_{g+1}^S(4,8)$ to one of the boundary component, say to V. Because of the theorem quoted above, this means that, scheme theoretically, the intersection of $Y_{\mathfrak{m}}$ and $V \cong \mathcal{A}_g(4,8)^S$ away from decomposable abelian varieties is isomorphic to $Y_{\mathfrak{n}}$; where $\mathfrak{n}$ is the g dimensional system of fundamental theta characteristic defined in equation 7 of [SM03]. $\square$

We now complete the description of the tangent space of Hyp^S_{g+1} along Hyp_g. Let C be a smooth genus g hyperelliptic curve and X its Jacobian. To start with, let us describe the normal bundle exact sequence of $\mathcal{A}_g$ in $\mathcal{A}^S_{g+1}$ at X. This sequence reads

$$0 \to T_X\mathcal{A}_g \to T_X\mathcal{A}^S_{g+1} \to H^0(X, 2\Theta)^\vee \to 0$$

We will need the following explicit description of the action of these derivations. Let F_{g+1} be a modular form on $\mathcal{A}_{g+1}$ and F_g its restriction to $\mathcal{A}_g$. For any element T in the Siegel upper half space $\mathcal{H}_{g+1}$ write

$$T = \begin{pmatrix} \tau & z \\ {}^tz & t \end{pmatrix},$$

with t in $\mathcal{H}_1$ and τ in $\mathcal{H}_g$. Let $q := \exp(2\pi i t)$; then the *Fourier-Jacobi* expansion of F_{g+1} is

$$F_{g+1}(T) = F_g(\tau) + \sum_{n\geq 1} f_n(\tau, z)q^n \tag{1}$$

where f_n is a section of $H^0(X_\tau, 2n\Theta)$, X_τ is the principally polarised abelian variety defined by τ, and z is a system of co-ordinates on X_τ. A derivation $D \in T_X\mathcal{A}_g$ acts as $D(F_{g+1}) = D(F_g)$; a derivation in $D \in H^0(X, 2\Theta)^\vee$ acts as $D(F_{g+1}) = D(f_1)$.

The normal bundle exact sequence for Hyp_g in Hyp^S_{g+1} at C is a subsequence of the normal bundle exact sequence of $\mathcal{A}_g$ in $\mathcal{A}^S_{g+1}$ at X. To describe it, we need to introduce the following morphism

$$\begin{array}{rcccccc} \Psi: & C & \xrightarrow{f} & C \times C & \xrightarrow{\delta} & X \\ & p & \mapsto & (p, \iota(p)) & & \\ & & & (a, b) & \mapsto & AJ(a) - AJ(b) \end{array} \tag{2}$$

where ι is the hyperelliptic involution and AJ is the Abel-Jacobi map.

Lemma 4.5. The pull-back $\Psi^*2\Theta$ is isomorphic to $2(K_C + W)$, where W is the divisor of Weierstrass points on C.

Proof. The pull-back $\delta^*2\Theta$ is $K_C \boxtimes K_C(2\Delta)$, where Δ is the diagonal; this is well-known, e.g. [Wel86, Equation 4.4]. Now, we pull-back $K_C \boxtimes K_C(2\Delta)$ via f. The pull-back of Δ is W; the pull-back of $K_C \boxtimes K_C$ is $K_C + \iota^*K_C = 2K_C$. □

Theorem 4.6. *Keep notation as above, the normal bundle exact sequence for Hyp_g in Hyp_{g+1} at C is*

$$0 \to T_C Hyp_g \to T_C Hyp^S_{g+1} \to P_C \to 0$$

where P_C is the image of the map

$$\Psi^*\colon H^0(X, 2\Theta) \to H^0(C, 2(K_C + W)).$$

In other words, the normal tangent cone at C is the cone over $(\Psi(C), 2\Theta)$.

Proof. There are two things we have to prove, first

$$T_C Hyp_{g+1}^S \cap T_C \mathcal{A}_g = T_C Hyp_g\,;$$

but this is equivalent to Theorem 4.3.
To describe the co-kernel of the inclusion

$$T_C Hyp_g \hookrightarrow T_C Hyp_{g+1}^S$$

we need to know that, after blowing up $\mathcal{A}_g^S$ in $\mathcal{A}_{g+1}^S$, the proper transform of Hyp_{g+1}^S meets the Kummer variety of X in $\Psi(C)$. This is proved in [Nam73, Theorem 6], just remark that to obtain a generic irreducible nodal hyperelliptic curve we need to glue two points conjugated under the hyperelliptic involution. □

In [Cod14, Lemma 3.6], it is shown that Ψ^* is not surjective and P_C has rank $2g$; however, we do not need this result here.

Remark 4.7 (Failure of Theorem 4.3 at the locus of decomposable abelian varieties). For the sake of completeness, let us sketch a proof of the failure of Theorem 4.3 at the locus of decomposable abelian varieties. This result is not needed in this paper, but we think that the study of this intersection is interesting on its own. Pick two integers such that $g_1 + g_2 = g$; fix a hyperelliptic curve C of genus g_1 and a hyperelliptic curve D of genus g_2. Call ι the hyperelliptic involution. The point (C, D) in Hyp_{g+1}^S represents all the hyperelliptic curves of the form $C \sqcup D/(p \sim q, \iota(p) \sim \iota(q))$, where p is a point varying in C and q is varying in D. Recall that we have an identification

$$\mathrm{Sym}^2(H^0(C, K_C)^\vee \oplus H^0(D, K_D)^\vee) = T_{(J(C)\times J(D))}\mathcal{A}_g$$

Arguing as in [CSB14], we can show that the tangent space of Hyp_{g+1}^S at (C, D) contains the image of the map

$$\begin{array}{cccc} \psi\colon & C/\iota \times D/\iota & \to & \mathbb{P}\,\mathrm{Sym}^2(H^0(C, K_C)^\vee \oplus H^0(D, K_D)^\vee) \\ & (p, q) & \mapsto & \omega_i(p)\psi_j(q) + \omega_j(p)\psi_i(q) \end{array}$$

where ω_i are a basis of $H^0(C, K_C)$ and ψ_i are a basis of $H^0(D, K_D)$. This is the same tangent direction we get when we consider an appropriate smoothing of the genus g nodal curve $C \sqcup D/p \sim q$; this smoothing is not hyperelliptic, so the intersection of $T_{(C,D)}Hyp_{g+1}^S$ with $T_{(J(C)\times J(D))}\mathcal{A}_g$ is strictly bigger than $T_{(C,D)}Hyp_g^S$.

5. Projective invariants of hyperelliptic curves

In this section, we review the theory of projective invariants, and we use it to show that certain modular forms vanish on the hyperelliptic locus.
To start with, let us introduce the auxiliary space $\mathcal{B}_g$. This is the moduli space of $2g + 2$ points on $\mathbb{P}^1$, up to permutation and projectivity. The points are counted with multiplicity, points are not allowed to have multiplicity bigger that $g + 1$. This space is classically constructed as a GIT quotient; it is irreducible,

it has an open dense subset $\mathcal{B}_g^\circ$ where the $2g+2$ points are all distinct, and a boundary D where at least two points coincide.
Let C be a smooth genus g hyperelliptic curve, fix a two to one map $\pi\colon C \to \mathbb{P}^1$. This morphism is unique up to projective transformations of $\mathbb{P}^1$, it ramifies at $2g+2$ points. A point p is called a Weierstrass point if it is a ramification point for π.

Definition 5.1 (Projective invariants). *The projective invariants of C are the image of the Weierstrass points under π, considered up to permutations and projective automorphisms of $\mathbb{P}^1$.*

Equivalently, the projective invariants of C are the points of the branch divisor of π, considered up to projectivity. The projective invariants of a smooth hyperelliptic curve C are naturally a point of $\mathcal{B}_g^\circ$, so we have a morphism

$$f_g\colon Hyp_g \to \mathcal{B}_g^\circ$$

One can reconstruct a hyperelliptic curve out of its projective invariants, and given $2g+2$ points on $\mathbb{P}^1$ there is a hyperelliptic curve with that projective invariants; this means that f_g is an isomorphism.
This construction has been extensively used to study the moduli space Hyp_g; references are [Igu67], [AL02] and [Pas11, Chapter 2].
As an aside, let us recall that the Thomæ's formula permits to write the cross-ratios of the projective invariants in term of second order theta functions evaluated at the period matrix of C.
Following [AL02], the map f_g extends to an isomorphism

$$f_g\colon Hyp_g \sqcup \eta_0^* \to \mathcal{B}_g^\circ \sqcup D^*$$

where η_0^* parametrises irreducible singular hyperelliptic curves with just one node, i.e. curves of the form $C/p \sim \iota(p)$, and D^* parametrises $2g+2$ points on $\mathbb{P}^1$ such that exactly 2 points coincide. The image of a curve in η_0^* is a set of $2g+2$ points of the form $\{p_1,\dots,p_{2g},p,p\}$, where $\{p_1,\dots,p_{2g}\}$ are the projective invariants of the normalisation and the glued points are the pre-images of p under π.
Always following [AL02], we can extend f_g further to a morphism

$$f_g\colon \overline{Hyp}_g \to \mathcal{B}_g$$

which contracts the boundary divisors of the Deligne-Mumford compactification different from the closure of η_0^* to high co-dimension loci.

We need some more information about $\mathcal{B}_g$, again references are [Igu67], [AL02] and [Pas11, Chapter 2]. From the GIT point of view, the moduli space $\mathcal{B}_g$ can be constructed as the Proj of the ring $S(2,2g+2)$, which is the ring of co-invariant of binary forms of degree $2g+2$. This ring is formally constructed as follows: let Z be the cartesian product of $2g+2$ copies of $\mathbb{P}^1$, on this variety we have a diagonal action of $SL(2,\mathbb{C})$ and an action of the symmetric group; this action linearise to an action on the line bundle $M := \mathcal{O}(1,\dots,1)$, the ring $S(2,2g+2)$ is the ring of invariant element of $\mathcal{R}(Z,M)$. More concretely,

$S(2, 2g+2)$ is the ring of symmetric functions in $2g+2$ variables, which are co-invariant under the natural action of $SL(2, \mathbb{C})$. The discriminant Δ is an element of $S(2, 2g+2)$ of degree $4g+2$, it cuts out the boundary divisor D.

We now consider also the Jacobian morphism

$$j\colon \overline{Hyp}_g \to Hyp_g^S \hookrightarrow \mathcal{A}_g$$

Take the composition

$$j \circ f_g^{-1}\colon \mathcal{B}_g \dashrightarrow \mathcal{A}_g$$

Under this map, D^* dominates Hyp_{g-1}; explicitly, the set $\{p_1, \ldots, p_{2g}, p, p\}$ is mapped to the Jacobian of the smooth hyperelliptic curve defined by the projective invariants $\{p_1, \ldots, p_{2g}\}$. Taking the pull-back we obtain a morphism of graded ring

$$\rho\colon \mathcal{R}(\mathcal{A}_g, L_g) \to S(2, 2g+2)$$

whose kernel is exactly the ideal of modular forms vanishing on the hyperelliptic locus. This map is sometime called Igusa morphism of projective invariants, it was introduced in [Igu67], where Igusa proved that its degree is $\frac{1}{2}g$ [1]. Using this construction we can prove the following criterion.

Criterion 5.2 (Weissauer - unpublished). *Let F_g be a weight n and degree g modular form. Restrict it to Hyp_g^S and say it vanishes along Hyp_{g-1} with multiplicity at least k. If*

$$\frac{n}{k} < 8 + \frac{4}{g}$$

then F_g vanishes on Hyp_g.

Proof. Suppose F_g vanishes with multiplicity at least k on Hyp_{g-1}. This means that $(j \circ f_g^{-1})^* F_g$ vanishes with multiplicity at least k on D. In other words, Δ^k divides $\rho(F_g)$. The degree of the discriminant in $S(2, 2g+2)$ is $4g+2$, the degree of $\rho(F_g)$ is $\frac{1}{2}gn$. Since, by hypothesis,

$$k(4g+2) > \frac{1}{2}gn$$

we obtain that $\rho(F_g)$ is equal to zero, so the claim. □

Remark 5.3 (Relation with Theorem 4.6). To show that F_g vanishes along Hyp_{g-1} with multiplicity at least 2 one needs to know the tangent space of Hyp_g^S along Hyp_{g-1}; in the applications, especially in Theorem 6.1, we will use the description given in Theorem 4.6.

Remark 5.4 (Other versions of Criterion 5.2). A weaker version of Criterion 5.2 can be found in [Poo96]. An alternative proof is in [Pas11]. In [SM00], Salvati Manni attributed this criterion to Weissauer, in an unpublished preprint, and showed that the inequality is sharp. This Criterion is also related to the slope of the hyperelliptic locus, in the sense of slope of the cone of effective divisors (cf. [CH88]).

[1] If g is odd, all non-trivial modular forms have odd degree, so the factor $\frac{1}{2}$ should not worry the reader.

Combining Theorem 1.3 and Criterion 5.2 we can find a first group of stable equations for the hyperelliptic locus. We will need the following basic invariant of a quadratic form (Q,Λ)

$$\mu(Q) := \min\{Q(v,v) \mid v \in \Lambda; v \neq 0\} = \min\{n \mid \mathcal{R}_n(\Lambda) \neq \varnothing\}\,,$$

where $\mathcal{R}_n(\Lambda)$ is the set of vectors of Λ of norm $2n$.

Theorem 5.5. *Let (Q,Λ) and (P,Γ) be two even positive definite unimodular quadratic forms of rank N and let $\mu := \min\{\mu(Q),\mu(P)\}$. If*

$$\frac{N}{\mu} \leq 8\,,$$

then

$$F := \Theta_Q - \Theta_P$$

is a stable equation for the hyperelliptic locus. In other words, F_g vanishes on Hyp_g for every g.

Proof. The proof is by induction on g. The difference of two theta series vanishes on $\mathcal{A}_0$. Suppose the statement true for g, we want to apply Criterion 5.2 to F_{g+1}. Call $k := \frac{1}{2}\mu$, we need to prove that F_{g+1} vanishes at the boundary component Hyp_g with multiplicity at least k.
We first compute the multiplicity along tangent direction parallel to the boundary, namely along $T_C Hyp^S_{g+1} \cap T_C\mathcal{A}_g$, where C is a smooth genus g hyperelliptic curve. This intersection is, by Theorem 1.3, equal to $T_C Hyp_g$. By induction, F_g vanishes along Hyp_g, so F_{g+1} is annihilated by the derivations contained in $T_C Hyp^S_{g+1} \cap T_C\mathcal{A}_g$.
Let us now look at the normal direction to $\mathcal{A}_g$; we will use the Fourier-Jacobi expansion introduced in the equation (1). Writing out the Fourier-Jacobi expansion of F_{g+1}, the hypothesis on μ implies that the first k terms vanish. This means that F_{g+1} vanishes with order at least k along the normal direction to $\mathcal{A}_g$ in $\mathcal{A}^S_{g+1}$; in particular, we obtain that it vanishes along Hyp_g with multiplicity at least k and we can apply Criterion 5.2. □

The hypotheses of Theorem 5.5 are quite restrictive; let us describe the cases where the Theorem can be applied.

Proposition 5.6. *Let (Q,Λ) be an even, positive definite, unimodular quadratic form, then*

$$\frac{\mathrm{rk}(Q)}{\mu(Q)} \leq 8$$

if and only if the pair $(\mathrm{rk}(Q),\mu(Q))$ is equal to one of the following pairs: $(8,2)$, $(16,2)$, $(32,4)$ or $(48,6)$.

Proof. Given any even unimodular quadratic form (Q,Λ), there is an upper bound

$$\mu(Q) \leq 2\lfloor\frac{\mathrm{rk}(Q)}{24}\rfloor + 2$$

where "$\lfloor\,\rfloor$" is the round down (see [CS99, Section 7.7 Corollary 21]). This bound, combined with the fact that the rank is divisible by 8, gives the Proposition.

□

Let us call the type of quadratic form the pair $(\mathrm{rk}(Q), \mu(Q))$. There is just one quadratic form of type $(8, 2)$ and one of type $(24, 4)$, so we do not get any stable equation in these cases. The rank 16 case was considered by Poor in [Poo96]: there are two quadratic forms of type $(16, 2)$, so one gets one equation. In [Kin03, Corollary 5], using a generalization of the mass formula, it is shown that there exist at least ten millions of quadratic forms of rank 32 and $\mu = 4$; however, just 15 of them are known explicitly. The situation for quadratic forms of type $(48, 6)$ is not clear: believably, there exist many of them, see [Kin03, Page 15], but there is not any lower bound and just 3 of them are known explicitly. (King adopts a slightly different notation: every quadratic form is tacitly assumed to be positive definite.) To summarise, Theorem 5.5 can be applied to the following cases

Corollary 5.7. *If P and Q are two even, unimodular, positive definite quadratic forms meeting one of the following two hypotheses:*

(1) $\mathrm{rk}(P) = \mathrm{rk}(Q) = 32$ *and* $\mu(P) = \mu(Q) = 4$*; that is, the quadratic forms do not have any vector of norm* 2*;*

(2) $\mathrm{rk}(P) = \mathrm{rk}(Q) = 48$ *and* $\mu(P) = \mu(Q) = 6$*; that is, the quadratic forms do not have any vector of norm* 2 *and* 4*;*

then, the difference

$$\Theta_P - \Theta_Q$$

is a stable equation for the hyperelliptic locus.

6. Niemeier quadratic forms

Niemeier quadratic forms are rank 24 quadratic forms. In this section we prove the following:

Theorem 6.1. *Let (P, Γ) and (Q, Λ) be two rank* 24 *quadratic forms with the same number of vectors of norm* 2*, then the difference*

$$\Theta_P - \Theta_Q$$

is a stable equation for the hyperelliptic locus.

Vectors of norm 2 are usually called roots. This result concerns the following 5 pairs of quadratic forms

quadratic forms	$A_5^4D_4$, D_4^6	$A_9^2D_6$, D_6^4	$A_{11}D_7E_6$, E_6^4
# roots	72	120	144

$A_{17}E_7$, $D_{10}E_7^2$	E_8D_{16}, E_8^3
216	360

where a quadratic form of rank 24 is labelled by its root system (see e.g. [Ebe13, Section 3] for more details). The pair $E_8^3, D_{16}E_8$ corresponds to the modular form $\Theta_{E_8}(\Theta_{E_8\oplus E_8} - \Theta_{D_{16}^+})$, so its behaviour was well-understood. The others cases can not be expressed as a product of lower weight stable modular forms and they are not covered by previous results.
This result is surprising because the slope of these quadratic forms, i.e. the ratio between the rank and the norm of the shortest vector, is strictly bigger than 8, so they were not expected to vanish on the hyperelliptic locus in every genus. Before proving our theorem we need two preliminary results.

6.1. A formula for sections of 2Θ. Let s be a section of 2Θ on the Jacobian of a curve C with period matrix τ. For every couple of points a and b of C the following classical formula holds:

$$s(\tau, a-b) = E(a,b)^2[s(\tau,0)\omega(a,b) + \sum_{i,j} \frac{\partial^2 s}{\partial z_i \partial z_j}(\tau,0)\omega_i(a)\omega_j(b)] \tag{3}$$

where E is the Prime form, $\{\omega_i\}$ is the basis of the holomorphic differentials on C corresponding to the basis $\{\frac{\partial}{\partial z_i}\}$ of the tangent space at the origin of the Jacobian, $\omega(a,b)$ is the fundamental normalised bi-differential, and everything is trivialised with respect to a choice of local co-ordinates z_a and z_b. This formula is well known, see e.g. [MV10, Appendix A].

6.2. The "heat equation" for Niemeier quadratic forms. The classification of rank 24 quadratic forms is due to Niemeier, but it has been simplified by Venkov proving and using the following identity

Theorem 6.2 (Venkov, cf. [Ebe13] Section 3). *Let (Λ, Q) be a rank 24 quadratic form, then*

$$r_2(\Lambda)Q(v,w) = 8 \sum_{y\in\mathcal{R}_2(\Lambda)} Q(y,w)Q(y,v) \qquad \forall\, v,w\in\Lambda$$

where $r_2(\Lambda)$ is the number of roots and $\mathcal{R}_2(\Lambda)$ is the set of roots.

The proof relies upon the theory of degree 1 modular forms with harmonic coefficients. Let us draw a consequence of this result about the Fourier-Jacobi expansion of theta series (cf. equation (1)).

Corollary 6.3 (Heat equation). *The first Fourier-Jacobi coefficient f_1 of a theta series associated to a rank 24 quadratic form Λ satisfies the following "heat equation"*

$$r_2(\Lambda)\pi i \frac{\partial f_1}{\partial \tau_{ij}}(\tau,0) = 3(1+\delta_{ij})\frac{\partial^2 f_1}{\partial z_i \partial z_j}(\tau,0)\,,$$

where $r_2(\Lambda)$ is the number of roots of Λ.

Proof. By explicit computation, we can write out the first Fourier-Jacobi coefficient of a theta series:

$$f_1(\tau,z)=\sum_{x_1,\dots,x_g\in\Lambda}\sum_{y\in\mathcal{R}_2(\Lambda)}\exp(\pi i\sum_{i,j}Q(x_i,x_j)\tau_{ij}+2\pi i\sum_i Q(y,x_i)z_i)\,.$$

Fix two indexes i and j, by explicit computation we have

$$\frac{\partial^2 f_1}{\partial z_i\partial z_j}(\tau,0)=(2\pi i)^2\sum_{x_1,\dots,x_g\in\Lambda}\sum_{y\in\mathcal{R}_2(\Lambda)}Q(y,x_i)Q(y,x_j)\exp(\pi i\sum_{i,j}Q(x_i,x_j)\tau_{ij})\,,$$

On the other hand

$$(1+\delta_{ij})\frac{\partial f_1}{\partial\tau_{ij}}(\tau,0)=2\pi i\sum_{x_1,\dots,x_g\in\Lambda}Q(x_i,x_j)\exp(\pi i\sum_{i,j}Q(x_i,x_j)\tau_{ij})\,,$$

the coefficient $(1+\delta_{ij})$ is because the variables on $\mathcal{H}_g$ are τ_{ij} with $i\le j$, so when we compute the derivative with respect to τ_{ij} we need to derive both τ_{ij} and τ_{ji}. Applying Theorem 6.2 we obtain the result. □

This formula is also discussed in [MV10, page 16]. This result is generalised to higher order Fourier-Jacobi coefficients and higher rank quadratic forms in [Cod14, Theorem 10.3].

6.3. Proof of Theorem 6.1. We want to show that $F_g=\Theta_{P,g}-\Theta_{Q,g}$ is zero on Hyp_g for every g; we argue by induction on g. The case $g=0$ is easy. To prove the inductive step we use Criterion 5.2: we need to show that F_{g+1} vanishes along Hyp_g with multiplicity at least 2. As in Theorem 5.5, the derivative along directions tangent to $\mathcal{A}_g$ vanishes because of Theorem 1.3 and the inductive hypothesis.

The normal direction is quite different: now the first Fourier-Jacobi coefficient is not trivial, so there is some work to do. Because of Theorem 4.6, it is enough to check that f_1 vanishes when restricted to points of the form $(\tau,p-\iota(p))$, where τ is the period matrix of a smooth hyperelliptic curve C, p is a point of C and ι is the hyperelliptic involution.

To show this we argue as follows. First remark that

$$f_1(\tau,0)=F_g(\tau)=0$$

Then we apply the formula (3), trivialising everything with respect to coordinates z_p and ι^*z_p and recalling that

$$\frac{\omega}{dz_p}(p)=\frac{\omega}{\iota^*dz_p}(\iota(p))$$

we get

$$f_1(\tau,p-\iota(p))=E(p,\iota(p))^2\sum_{i,j}\frac{\partial^2 f_1}{\partial z_i\partial z_j}(\tau,0)\omega_i(p)\omega_j(p)$$

Now the heat equation 6.3 and the hypothesis on the number of roots come into the game: since $r_2(\Lambda) = r_2(\Gamma) =: r$, we have

$$6\sum_{i,j}\frac{\partial^2 f_1}{\partial z_i \partial z_j}(\tau,0)\omega_i(p)\omega_j(p) = r\pi i\sum_{i\geq j}\frac{\partial F_g}{\partial \tau_{ij}}(\tau)\omega_i(p)\omega_j(p) = (r\pi i)dF_g(\tau)(p)$$

Let us explain the last equality: the fibre of the cotangent bundle of $\mathcal{A}_g$ at $Jac(C)$ is isomorphic to $\operatorname{Sym}^2 H^0(C,K_C)$, so $dF_g(\tau)$ is a quadric in $\mathbb{P}H^0(C,K_C)^\vee$ and we can evaluate it on the image of p under the canonical map.
The co-normal bundle of Hyp_g in $\mathcal{M}_g$ is given by the -1 eigenspace of $H^0(C,2K_C)$; the image of the co-differential $m\colon \operatorname{Sym}^2 H^0(C,K_C) \to H^0(C,2K_C)$ is the $+1$ eigenspace; we conclude that the quadric in the co-normal bundle of Hyp_g in $\mathcal{A}_g$ vanishes along the canonical image of C. Since F_g vanishes along Hyp_g, dF_g is a quadric containing the canonical image of C; in other words, it has to vanish when evaluated at any point p of C. This concludes the proof of Theorem 6.1.

6.4. Other results about Niemeir quadratic forms. With similar tools, we can prove other results about the behaviour of these modular forms on the moduli space of curves and abelian varieties.

Theorem 6.4 ([Cod14] Corollary 11.2). *Let P and Q be two even positive definite unimodular quadratic forms of rank 24 with the same number of roots, then the stable modular form*

$$F := \Theta_P - \Theta_Q$$

is zero on $\mathcal{M}_g$ for $g \leq 4$, and it cuts a divisor of slope 12 on $\mathcal{M}_5$.

Theorem 6.5 ([Cod14] Theorem 11.3). *The following degree 5 modular forms are non-trivial cusp forms*

$$\Theta(D_{16}E_8) - \Theta(E_8^3) - \frac{21504}{24}(\Theta(A_5^4 D_5) - \Theta(D_4^6))$$
$$\Theta(D_{16}E_8) - \Theta(E_8^3) - \frac{21504}{216}(\Theta(A_9^2 D_6) - \Theta(D_6^4))$$
$$\Theta(D_{16}E_8) - \Theta(E_8^3) - \frac{21504}{480}(\Theta(A_{11}D_7E_6) - \Theta(E_6^4))$$
$$\Theta(D_{16}E_8) - \Theta(E_8^3) - \frac{21504}{-2520}(\Theta(A_{17}E_7) - \Theta(D_{10}E_7^2))$$

where, for typographical reasons, we write $\Theta(Q)$ rather than $\Theta_{Q,5}$.

References

[ACG11] E. Arbarello, M. Cornalba, and P. A. Griffiths. *Geometry of algebraic curves. Volume II*, volume 268 of *Grundlehren der Mathematischen Wissenschaften*. Springer, Heidelberg, 2011. With a contribution by J. D. Harris.

[AL02] D. Avritzer and H. Lange. The moduli spaces of hyperelliptic curves and binary forms. *Mathematische Zeitschrift*, 242(4):615–632, 2002.

[BvdGHZ08] J. H. Bruinier, G. van der Geer, G. Harder, and D. Zagier. *The 1-2-3 of modular forms*. Universitext. Springer-Verlag, Berlin, 2008.

[CH88] M. Cornalba and J. Harris. Divisor classes associated to families of stable varieties, with applications to the moduli space of curves. *Ann. Sci. École Norm. Sup. (4)*, 21(3):455–475, 1988.

[Cod14] G. Codogni. *Satake compactifications, Lattices and Schottky problem*. PhD thesis, University of Cambridge, 2014.

[CS99] J. H. Conway and N. J. A. Sloane. *Sphere packings, lattices and groups*, volume 290 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, New York, 1999.

[CSB14] G. Codogni and N. I. Shepherd-Barron. The non-existence of stable Schottky forms. *Compos. Math.*, 150(4):679–690, 2014.

[Ebe13] W. Ebeling. *Lattices and codes*. Advanced Lectures in Mathematics. Springer Spektrum, Wiesbaden, third edition, 2013. A course partially based on lectures by Friedrich Hirzebruch.

[Fre77] E. Freitag. Stabile Modulformen. *Math. Ann.*, 230(3):197–211, 1977.

[Fre83] E. Freitag. *Siegelsche Modulfunktionen*, volume 254 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, 1983.

[GHT14] S. Grushevsky, K. Hulek, and O. Tommasi. Stable cohomology of the perfect cone toroidal compactification of the moduli space of abelian varieties. 2014. To appear in *Journal für die reine und angewandte Mathematik*.

[GKV10] M. R. Gaberdiel, C. A. Keller, and R. Volpato. Genus two partition functions of chiral conformal field theories. *Commun.Num.Theor.Phys*, (4):295–364, 2010.

[GV09] M. R. Gaberdiel and R. Volpato. Higher genus partition functions of meromorphic conformal field theories. *J. High Energy Phys.*, (6):048, 43, 2009.

[Hoy63] W. L. Hoyt. On products of algebraic families of jacobian varieties. *Annals of Mathematics*, 77(3):415–423, 1963.

[Igu67] J. Igusa. Modular forms and projective invariants. *Amer. J. Math.*, 89:817–855, 1967.

[Igu81] J. Igusa. Schottky's invariant and quadratic forms. In *E. B. Christoffel (Aachen/Monschau, 1979)*, pages 352–362. Birkhäuser, Basel, 1981.

[Kin03] O. D. King. A mass formula for unimodular lattices with no roots. *Mathematics of Computation*, 72(242):839–863 (electronic), 2003.

[Kum02] S. Kumar. *Kac-Moody groups, their flag varieties and representation theory*, volume 204 of *Progress in Mathematics*. Birkhäuser Boston Inc., 2002.

[Mat15] M. Matone. A surprising relation for the effective coupling constants of N=2 super yang-mills theories. *J. High Energy Phys.*, 1506 (2015), 109.

[Mum07] D. Mumford. *Tata lectures on theta. I.* Modern Birkhäuser Classics. Birkhäuser Boston Inc., Boston, MA, 2007. With the collaboration of C. Musili, M. Nori, E. Previato and M. Stillman, Reprint of the 1983 edition.

[MV10] M. Matone and R. Volpato. Getting superstring amplitudes by degenerating Riemann surfaces. *Nuclear Phys. B*, 839(1-2):21–51, 2010.

[Nam73] Y. Namikawa. On the canonical holomorphic map from the moduli space of stable curves to the Igusa monoidal transform. *Nagoya Math. J.*, 52:197–259, 1973.

[Pas11] S. Pascolutti. *Boundaries of Moduli Space of Hyperelliptic Curves with a Level 2 Structure*. PhD thesis, La Sapienza, Università di Roma, 2011.

[Poo96] C. Poor. Schottky's form and the hyperelliptic locus. *Proc. Amer. Math. Soc.*, 124(7):1987–1991, 1996.

[Sat56] I. Satake. On the compactification of the Siegel space. *J. Indian Math. Soc. (N.S.)*, 20:259–281, 1956.

[SB13] N. Shepherd-Barron. Siegel modular forms and the gonality of curves. *arXiv:1306.6253*, 2013.

[SM89] R. Salvati Manni. Thetanullwerte and stable modular forms. *Amer. J. Math.*, 111 (3) 435–455, 1989

[SM00] R. Salvati Manni. Slope of cusp forms and theta series. *J. Number Theory*, 83 (2) 282–296, 2000

[SM03] R. Salvati Manni. Modular forms vanishing on the hyperelliptic locus. *Japan. J. Math. (N.S.)*, 29 (1) 135–142, 2003

[Tsu91] S. Tsuyumine. Thetanullwerte on a moduli space of curves and hyperelliptic loci. *Mathematische Zeitschrift*, 207 (4) :539–568, 1991.

[Wel86] G. E. Welters. The surface $C - C$ on Jacobi varieties and 2nd order theta functions. *Acta Math.*, 157(1-2):1–22, 1986.

Giulio Codogni
Dipartimento
di Matematica e Fisica
Università degli
Studi Roma Tre
Largo S. Leonardo Murialdo 1
00146 Rome (Italy)
codogni@mat.uniroma3.it

Some Results on Bessel Functionals for GSp(4)

Brooks Roberts and Ralf Schmidt[1]

Received: January 5, 2015
Revised: December 12, 2015

Communicated by Don Blasius

Abstract. We prove that every irreducible, admissible representation π of $\mathrm{GSp}(4, F)$, where F is a non-archimedean local field of characteristic zero, admits a Bessel functional, provided π is not one-dimensional. If π is not supercuspidal, we explicitly determine the set of all Bessel functionals admitted by π, and prove that Bessel functionals of a fixed type are unique. If π is supercuspidal, we do the same for all split Bessel functionals.

2010 Mathematics Subject Classification: Primary 11F70 and 22E50

Contents

[1] Supported by NSF grant DMS-1100541.

Introduction

The uniqueness and existence of appropriate *models* for irreducible, admissible representations of a linear reductive group over a local field has long played an important role in local and global representation theory. Best known are perhaps the Whittaker models for general linear groups, which are instrumental in proving multiplicity one theorems and the analytic properties of automorphic L-functions. Generic representations, i.e., those admitting a Whittaker model, have an important place in the representation theory of $\mathrm{GSp}(4)$ as well, the group under consideration in this paper; see [16] for an early example of their use. For $\mathrm{GSp}(4)$ they play a less comprehensive role, however, since there are many important non-generic automorphic representations, for example those generated by holomorphic Siegel modular forms.

The use of *Bessel models*, or equivalently *Bessel functionals*, as a substitute for the often missing Whittaker models for GSp(4) has been pioneered by Novodvorsky and Piatetski-Shapiro. Similar to the generic case, Bessel models consist of functions on the group with a simple transformation property under a certain subgroup; see below for precise definitions. The papers [17] and [15] are concerned with the uniqueness of Bessel functionals in the case of trivial central character; the first paper treats the case of so-called *special* Bessel functionals. For the use of Bessel models in the study of analytic properties and special values of L-functions for non-generic representations, see [19], [36], [6], [21].
In this paper, we further investigate the existence and uniqueness of Bessel functionals for irreducible, admissible representations of $\mathrm{GSp}(4,F)$, where F is a non-archimedean local field of characteristic zero. To explain our results, we have to introduce some notation. Let F be a non-archimedean local field of characteristic zero, and let ψ be a non-trivial character of F. Let $\mathrm{GSp}(4,F)$ be the subgroup of g in $\mathrm{GL}(4,F)$ satisfying ${}^tgJg=\lambda(g)J$ for some scalar $\lambda(g)$ in $F^\times$, where

$$J=\begin{bmatrix}&&&1\\&&1&\\&-1&&\\-1&&&\end{bmatrix}.$$

The Siegel parabolic subgroup P of $\mathrm{GSp}(4,F)$ is the subgroup consisting of all matrices whose lower left 2×2 block is zero. Let N be the unipotent radical of P. The characters θ of N are in one-to-one correspondence with symmetric 2×2 matrices S over F via the formula

$$\theta(\left[\begin{smallmatrix}1&X\\&1\end{smallmatrix}\right])=\psi(\mathrm{tr}(S\left[\begin{smallmatrix}&1\\1&\end{smallmatrix}\right]X)).$$

We say that θ is *non-degenerate* if the matrix S is invertible, and we say that θ is *split* if $\mathrm{disc}(S)=1$; here $\mathrm{disc}(S)$ is the class of $-\det(S)$ in $F^\times/F^{\times 2}$. For a fixed S, we define

$$T=\left[\begin{smallmatrix}&1\\1&\end{smallmatrix}\right]\{g\in\mathrm{GL}(2,F):\,{}^tgSg=\det(g)S\}\left[\begin{smallmatrix}&1\\1&\end{smallmatrix}\right]. \tag{1}$$

We embed T into $\mathrm{GSp}(4,F)$ via the map

$$t\mapsto\begin{bmatrix}t&\\&\det(t)t'\end{bmatrix},$$

where for a 2×2-matrix g we write $g'=\left[\begin{smallmatrix}&1\\1&\end{smallmatrix}\right]{}^tg^{-1}\left[\begin{smallmatrix}&1\\1&\end{smallmatrix}\right]$. The group T normalizes N, so that we can define the semidirect product $D=TN$. This will be referred to as the *Bessel subgroup* corresponding to S. For t in T and n in N, we have $\theta(tnt^{-1})=\theta(n)$. Thus, if Λ is a character of T, we can define a character $\Lambda\otimes\theta$ of D by $(\Lambda\otimes\theta)(tn)=\Lambda(t)\theta(n)$. Whenever we regard $\mathbb{C}$ as a one-dimensional representation of D via this character, we denote it by $\mathbb{C}_{\Lambda\otimes\theta}$. Let (π,V) be an irreducible, admissible representation of $\mathrm{GSp}(4,F)$. A non-zero element of the space $\mathrm{Hom}_D(\pi,\mathbb{C}_{\Lambda\otimes\theta})$ is called a (Λ,θ)-*Bessel functional* for π. We say that π admits a (Λ,θ)-Bessel functional if $\mathrm{Hom}_D(\pi,\mathbb{C}_{\Lambda\otimes\theta})$ is non-zero, and that π admits a unique (Λ,θ)-Bessel functional if $\mathrm{Hom}_D(\pi,\mathbb{C}_{\Lambda\otimes\theta})$ is one-dimensional.

In this paper we prove three main results about irreducible, admissible representations π of $\mathrm{GSp}(4,F)$:

- If π is not one-dimensional, we prove that π admits some (Λ,θ)-Bessel functional; see Theorem 6.1.4.
- If θ is split, we determine the set of Λ for which π admits a (Λ,θ)-Bessel functional, and prove that such functionals are unique; see Proposition 3.4.2, Theorem 6.1.4, Theorem 6.2.2 and Theorem 6.3.2.
- If π is non-supercuspidal, or is in an L-packet with a non-supercuspidal representation, we determine the set of (Λ,θ) for which π admits a (Λ,θ)-Bessel functional, and prove that such functionals are unique; see Theorem 6.2.2 and Theorem 6.3.2.

We point out that all our results hold independently of the residual characteristic of F.

To investigate (Λ,θ)-Bessel functionals for (π,V) we use the P_3-module V_{Z^J}, the G^J-module $V_{Z^J,\psi}$, and the twisted Jacquet module $V_{N,\theta}$. Here,

$$P_3=\mathrm{GL}(3,F)\cap\begin{bmatrix}*&*&*\\ *&*&*\\ &&1\end{bmatrix},\quad Z^J=\mathrm{GSp}(4,F)\cap\begin{bmatrix}1&&&*\\ &1&&\\ &&1&\\ &&&1\end{bmatrix},$$
$$\text{and}\ \ G^J=\mathrm{GSp}(4,F)\cap\begin{bmatrix}1&*&*&*\\ &*&*&*\\ &*&*&*\\ &&&1\end{bmatrix}.$$

The P_3-module V_{Z^J} was computed for all π with trivial central character in [28]; in this paper, we note that these results extend to the general case. The G^J-module $V_{Z^J,\psi}$ is closely related to representations of the metaplectic group $\widetilde{\mathrm{SL}}(2,F)$. The twisted Jacquet module $V_{N,\theta}$ is especially relevant for non-supercuspidal representations. Indeed, we completely calculate twisted Jacquet modules of representations parabolically induced from the Klingen or Siegel parabolic subgroups. These methods suffice to treat most representations; for the few remaining families of representations we use theta lifts. As a by-product of our investigations we obtain a characterization of non-generic representations. Namely, the following conditions are equivalent: π is non-generic; the twisted Jacquet module $V_{N,\theta}$ is finite-dimensional for all non-degenerate θ; the twisted Jacquet module $V_{N,\theta}$ is finite-dimensional for all split θ; the G^J-module $V_{Z^J,\psi}$ is of finite length. See Theorem 7.1.4.

If an irreducible, admissible representation π admits a (Λ,θ)-Bessel functional, then π has an associated Bessel model. For unramified π admitting a (Λ,θ)-Bessel functional, the works [36] and [4] contain explicit formulas for the spherical vector in such a Bessel model. Other explicit formulas in certain cases of Iwahori-spherical representations appear in [32], [20] and [22]. We note that these works show that all the values of a certain vector in the given Bessel model can be expressed in terms of data depending only on the representation and Λ and θ; in this situation it follows that the Bessel functional is unique. As far as we know, a detailed proof of uniqueness of Bessel functionals in all cases has not yet appeared in the literature.

In the case of odd residual characteristic, and when π appears in a generic L-packet, the main local theorem of [23] gives an ε-factor criterion for the existence of a (Λ,θ)-Bessel functional. There is some overlap between the methods of [23] and the present work. However, the goal of this work is to give a complete and ready account of Bessel functionals for all non-supercuspidal representations. We hope these results will be useful for applications where such specific knowledge is needed.

1 Some definitions

Throughout this work let F be a non-archimedean local field of characteristic zero. Let $\bar{F}$ be a fixed algebraic closure of F. We fix a non-trivial character $\psi:\ F\to\mathbb{C}^\times$. The symbol $\mathfrak{o}$ denotes the ring of integers of F, and $\mathfrak{p}$ is the maximal ideal of $\mathfrak{o}$. We let ϖ be a fixed generator of $\mathfrak{p}$. We denote by $|\cdot|$ the normalized absolute value on F, and by ν its restriction to $F^\times$. The Hilbert symbol of F will be denoted by $(\cdot,\cdot)_F$. If Λ is a character of a group, we denote by $\mathbb{C}_\Lambda$ the space of the one-dimensional representation whose action is given by Λ. If $x=\left[\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right]$ is a 2×2 matrix, then we set $x^*=\left[\begin{smallmatrix}d&-b\\-c&a\end{smallmatrix}\right]$. If X is an l-space, as in 1.1 of [3], and V is a complex vector space, then $\mathcal{S}(X,V)$ is the space of locally constant functions $X\to V$ with compact support. Let G be an l-group, as in [3], and let H be a closed subgroup. If ρ is a smooth representation of H, we define the compactly induced representation (unnormalized) $\text{c-Ind}_H^G(\rho)$ as in 2.22 of [3]. If (π,V) is a smooth representation of G, and if θ is a character of H, we define the twisted Jacquet module $V_{H,\theta}$ as the quotient $V/V(H,\theta)$, where $V(H,\theta)$ is the span of all vectors $\pi(h)v-\theta(h)v$ for all h in H and v in V.

1.1 Groups

Let

$$\mathrm{GSp}(4,F)=\{g\in\mathrm{GL}(4,F):\ {}^tgJg=\lambda(g)J,\ \lambda(g)\in F^\times\},\qquad J=\begin{bmatrix}&&&1\\&&1&\\&-1&&\\-1&&&\end{bmatrix}.$$

The scalar $\lambda(g)$ is called the *multiplier* or *similitude factor* of the matrix g. The *Siegel parabolic subgroup* P of $\mathrm{GSp}(4,F)$ consists of all matrices whose lower left 2×2 block is zero. For a matrix $A\in\mathrm{GL}(2,F)$ set $A'=\left[\begin{smallmatrix}&1\\1&\end{smallmatrix}\right]{}^tA^{-1}\left[\begin{smallmatrix}&1\\1&\end{smallmatrix}\right]$. Then the Levi decomposition of P is $P=MN$, where

$$M=\{\left[\begin{smallmatrix}A&\\&\lambda A'\end{smallmatrix}\right]:\ A\in\mathrm{GL}(2,F),\ \lambda\in F^\times\},\tag{2}$$

and

$$N=\{\begin{bmatrix}1&&y&z\\&1&x&y\\&&1&\\&&&1\end{bmatrix}:\ x,y,z\in F\}.\tag{3}$$

Let Q be the *Klingen parabolic subgroup*, i.e.,

$$Q=\mathrm{GSp}(4,F)\cap\begin{bmatrix}*&*&*&*\\&*&*&*\\&*&*&*\\&&&*\end{bmatrix}.\tag{4}$$

The Levi decomposition for Q is $Q = M_Q N_Q$, where

$$M_Q = \{ \begin{bmatrix} t&&\\&A&\\&&t^{-1}\det(A)\end{bmatrix} :\ A \in \mathrm{GL}(2,F),\ t \in F^\times \}, \tag{5}$$

and N_Q is the *Heisenberg group*

$$N_Q = \{ \begin{bmatrix} 1&x&y&z\\&1&&y\\&&1&-x\\&&&1\end{bmatrix} :\ x,y,z \in F \}. \tag{6}$$

The subgroup of Q consisting of all elements with $t = 1$ and $\det(A) = 1$ is called the *Jacobi group* and is denoted by G^J. The center of G^J is $Z^J = \begin{bmatrix} 1&&&*\\&1&&\\&&1&\\&&&1\end{bmatrix}$. The standard Borel subgroup of $\mathrm{GSp}(4,F)$ consists of all upper triangular matrices in $\mathrm{GSp}(4,F)$. We let

$$U = \mathrm{GSp}(4,F) \cap \begin{bmatrix} 1&*&*&*\\&1&*&*\\&&1&*\\&&&1\end{bmatrix}$$

be its unipotent radical. The following elements of $\mathrm{GSp}(4,F)$ represent generators for the eight-element Weyl group,

$$s_1 = \begin{bmatrix} &1&&\\1&&&\\&&&1\\&&1&\end{bmatrix} \qquad \text{and} \qquad s_2 = \begin{bmatrix} 1&&&\\&&1&\\&-1&&\\&&&1\end{bmatrix}. \tag{7}$$

1.2 Representations

For a smooth representation π of $\mathrm{GSp}(4,F)$ or $\mathrm{GL}(2,F)$, we denote by $\pi^\vee$ its smooth contragredient.
For c_1, c_2 in $F^\times$, let ψ_{c_1,c_2} be the character of U defined by

$$\psi_{c_1,c_2}(\begin{bmatrix} 1&x&*&*\\&1&y&*\\&&1&-x\\&&&1\end{bmatrix}) = \psi(c_1 x + c_2 y). \tag{8}$$

An irreducible, admissible representation (π, V) of $\mathrm{GSp}(4,F)$ is called *generic* if the space $\mathrm{Hom}_U(V, \psi_{c_1,c_2})$ is non-zero. This definition is independent of the choice of c_1, c_2. It is known by [30] that, if non-zero, the space $\mathrm{Hom}_U(V, \psi_{c_1,c_2})$ is one-dimensional. Hence, π can be realized in a unique way as a space of functions $W :\ \mathrm{GSp}(4,F) \to \mathbb{C}$ with the transformation property

$$W(ug) = \psi_{c_1,c_2}(u) W(g), \qquad u \in U,\ g \in \mathrm{GSp}(4,F),$$

on which π acts by right translations. We denote this model of π by $\mathcal{W}(\pi, \psi_{c_1,c_2})$, and call it the *Whittaker model* of π with respect to c_1, c_2.
We will employ the notation of [35] for parabolically induced representations of $\mathrm{GSp}(4,F)$ (all parabolic induction is normalized). For details we refer to the summary given in Sect. 2.2 of [28]. Let χ_1, χ_2 and σ be characters of $F^\times$. Then

$\chi_1 \times \chi_2 \rtimes \sigma$ denotes the representation of $\mathrm{GSp}(4,F)$ parabolically induced from the character of the Borel subgroup which is trivial on U and is given by

$$\mathrm{diag}(a,b,cb^{-1},ca^{-1}) \longmapsto \chi_1(a)\chi_2(b)\sigma(c), \qquad a,b,c \in F^\times,$$

on diagonal elements. Let σ be a character of $F^\times$ and π be an admissible representation of $\mathrm{GL}(2,F)$. Then $\pi \rtimes \sigma$ denotes the representation of $\mathrm{GSp}(4,F)$ parabolically induced from the representation

$$\left[\begin{smallmatrix} A & * \\ & cA' \end{smallmatrix}\right] \longmapsto \sigma(c)\pi(A), \qquad A \in \mathrm{GL}(2,F),\ c \in F^\times, \tag{9}$$

of the Siegel parabolic subgroup P. Let χ be a character of $F^\times$ and π an admissible representation of $\mathrm{GSp}(2,F) \cong \mathrm{GL}(2,F)$. Then $\chi \rtimes \pi$ denotes the representation of $\mathrm{GSp}(4,F)$ parabolically induced from the representation

$$\left[\begin{smallmatrix} t & * & * \\ & g & * \\ & & \det(g)t^{-1} \end{smallmatrix}\right] \longmapsto \chi(t)\pi(g), \qquad t \in F^\times,\ g \in \mathrm{GL}(2,F), \tag{10}$$

of the Klingen parabolic subgroup Q.

For a character ξ of $F^\times$ and a representation (π, V) of $\mathrm{GSp}(4,F)$, the *twist* $\xi\pi$ is the representation of $\mathrm{GSp}(4,F)$ on the same space V given by $(\xi\pi)(g) = \xi(\lambda(g))\pi(g)$ for g in $\mathrm{GSp}(4,F)$, where λ is the multiplier homomorphism defined above. A similar definition applies to representations π of $\mathrm{GL}(2,F)$; in this case, the multiplier is replaced by the determinant. The behavior of parabolically induced representations under twisting is as follows,

$$\begin{aligned} \xi(\chi_1 \times \chi_2 \rtimes \sigma) &= \chi_1 \times \chi_2 \rtimes \xi\sigma, \\ \xi(\pi \rtimes \sigma) &= \pi \rtimes \xi\sigma, \\ \xi(\chi \rtimes \pi) &= \chi \rtimes \xi\pi. \end{aligned}$$

The irreducible constituents of *all* parabolically induced representations of $\mathrm{GSp}(4,F)$ have been determined in [35]. The following table, which is essentially a reproduction of Table A.1 of [28], provides a summary of these irreducible constituents. In the table, $\chi, \chi_1, \chi_2, \xi$ and σ stand for characters of $F^\times$; the symbol ν denotes the normalized absolute value; π stands for an irreducible, admissible, supercuspidal representation of $\mathrm{GL}(2,F)$, and ω_π denotes the central character of π. The trivial character of $F^\times$ is denoted by $1_{F^\times}$, the trivial representation of $\mathrm{GL}(2,F)$ by $1_{\mathrm{GL}(2)}$ or $1_{\mathrm{GSp}(2)}$, depending on the context, the trivial representation of $\mathrm{GSp}(4,F)$ by $1_{\mathrm{GSp}(4)}$, the Steinberg representation of $\mathrm{GL}(2,F)$ by $\mathrm{St}_{\mathrm{GL}(2)}$ or $\mathrm{St}_{\mathrm{GSp}(2)}$, depending on the context, and the Steinberg representation of $\mathrm{GSp}(4,F)$ by $\mathrm{St}_{\mathrm{GSp}(4)}$. The names of the representations given in the "representation" column are taken from [35]. The "tempered" column indicates the condition on the inducing data under which a representation is tempered. The "L^2" column indicates which representations are square integrable after an appropriate twist. Finally, the "g" column indicates which representations are generic.

	constituents of		representation	tempered	L^2	g
I	$\chi_1 \times \chi_2 \rtimes \sigma$		(irreducible)	χ_i, σ unitary		$\bullet$
II	$\nu^{1/2}\chi \times \nu^{-1/2}\chi \rtimes \sigma$	a	$\chi\mathrm{St}_{\mathrm{GL}(2)} \rtimes \sigma$	χ, σ unitary		$\bullet$
	$(\chi^2 \neq \nu^{\pm 1}, \chi \neq \nu^{\pm 3/2})$	b	$\chi 1_{\mathrm{GL}(2)} \rtimes \sigma$			
III	$\chi \times \nu \rtimes \nu^{-1/2}\sigma$	a	$\chi \rtimes \sigma\mathrm{St}_{\mathrm{GSp}(2)}$	χ, σ unitary		$\bullet$
	$(\chi \notin \{1, \nu^{\pm 2}\})$	b	$\chi \rtimes \sigma 1_{\mathrm{GSp}(2)}$			
IV	$\nu^2 \times \nu \rtimes \nu^{-3/2}\sigma$	a	$\sigma\mathrm{St}_{\mathrm{GSp}(4)}$	σ unitary	$\bullet$	$\bullet$
		b	$L(\nu^2, \nu^{-1}\sigma\mathrm{St}_{\mathrm{GSp}(2)})$			
		c	$L(\nu^{3/2}\mathrm{St}_{\mathrm{GL}(2)}, \nu^{-3/2}\sigma)$			
		d	$\sigma 1_{\mathrm{GSp}(4)}$			
V	$\nu\xi \times \xi \rtimes \nu^{-1/2}\sigma$	a	$\delta([\xi, \nu\xi], \nu^{-1/2}\sigma)$	σ unitary	$\bullet$	$\bullet$
	$(\xi^2 = 1,\ \xi \neq 1)$	b	$L(\nu^{1/2}\xi\mathrm{St}_{\mathrm{GL}(2)}, \nu^{-1/2}\sigma)$			
		c	$L(\nu^{1/2}\xi\mathrm{St}_{\mathrm{GL}(2)}, \xi\nu^{-1/2}\sigma)$			
		d	$L(\nu\xi, \xi \rtimes \nu^{-1/2}\sigma)$			
VI	$\nu \times 1_{F^\times} \rtimes \nu^{-1/2}\sigma$	a	$\tau(S, \nu^{-1/2}\sigma)$	σ unitary		$\bullet$
		b	$\tau(T, \nu^{-1/2}\sigma)$	σ unitary		
		c	$L(\nu^{1/2}\mathrm{St}_{\mathrm{GL}(2)}, \nu^{-1/2}\sigma)$			
		d	$L(\nu, 1_{F^\times} \rtimes \nu^{-1/2}\sigma)$			
VII	$\chi \rtimes \pi$		(irreducible)	χ, π unitary		$\bullet$
VIII	$1_{F^\times} \rtimes \pi$	a	$\tau(S, \pi)$	π unitary		$\bullet$
		b	$\tau(T, \pi)$	π unitary		
IX	$\nu\xi \rtimes \nu^{-1/2}\pi$	a	$\delta(\nu\xi, \nu^{-1/2}\pi)$	π unitary	$\bullet$	$\bullet$
	$(\xi \neq 1,\ \xi\pi = \pi)$	b	$L(\nu\xi, \nu^{-1/2}\pi)$			
X	$\pi \rtimes \sigma$		(irreducible)	π, σ unitary		$\bullet$
XI	$\nu^{1/2}\pi \rtimes \nu^{-1/2}\sigma$	a	$\delta(\nu^{1/2}\pi, \nu^{-1/2}\sigma)$	π, σ unitary	$\bullet$	$\bullet$
	$(\omega_\pi = 1)$	b	$L(\nu^{1/2}\pi, \nu^{-1/2}\sigma)$			
Va*	(supercuspidal)		$\delta^*([\xi, \nu\xi], \nu^{-1/2}\sigma)$	σ unitary	$\bullet$	
XIa*	(supercuspidal)		$\delta^*(\nu^{1/2}\pi, \nu^{-1/2}\sigma)$	π, σ unitary	$\bullet$	

In addition to all irreducible, admissible, non-supercuspidal representations, the table also includes two classes of supercuspidal representations denoted by Va* and XIa*. The reason that these supercuspidal representations are included in the table is that they are in L-packets with some non-supercuspidal representations. Namely, the Va representation $\delta([\xi,\nu\xi],\nu^{-1/2}\sigma)$ and the Va* representation $\delta^*([\xi,\nu\xi],\nu^{-1/2}\sigma)$ form an L-packet, and the XIa representation $\delta(\nu^{1/2}\pi,\nu^{-1/2}\sigma)$ and the XIa* representation $\delta^*(\nu^{1/2}\pi,\nu^{-1/2}\sigma)$ form an L-packet; see the paper [8]. Incidentally, the other non-singleton L-packets involving non-supercuspidal representations are the two-element packets $\{\tau(S,\nu^{-1/2}\sigma),\tau(T,\nu^{-1/2}\sigma)\}$ (type VIa and VIb), as well as $\{\tau(S,\pi),\tau(T,\pi)\}$ (type VIIIa and VIIIb).

2 Generalities on Bessel functionals

In this section we gather some definitions, notation, and basic results about Bessel functionals.

2.1 Quadratic extensions

Let $D\in F^\times$. If $D\notin F^{\times 2}$, then let $\Delta=\sqrt{D}$ be a square root of D in $\bar{F}$, and $L=F(\Delta)$. If $D\in F^{\times 2}$, then let $\sqrt{D}$ be a square root of D in $F^\times$, $L=F\times F$, and $\Delta=(-\sqrt{D},\sqrt{D})\in L$. In both cases L is a two-dimensional F-algebra containing F, $L=F+F\Delta$, and $\Delta^2=D$. We will abuse terminology slightly, and refer to L as the *quadratic extension associated to D*. We define a map $\gamma:L\to L$ called *Galois conjugation* by $\gamma(x+y\Delta)=x-y\Delta$. Then $\gamma(xy)=\gamma(x)\gamma(y)$ and $\gamma(x+y)=\gamma(x)+\gamma(y)$ for $x,y\in L$, and the fixed points of γ are the elements of F. The group $\mathrm{Gal}(L/F)$ of F-automorphisms $\alpha:L\to L$ is $\{1,\gamma\}$. We define norm and trace functions $\mathrm{N}_{L/F}:L\to F$ and $\mathrm{T}_{L/F}:L\to F$ by $\mathrm{N}_{L/F}(x)=x\gamma(x)$ and $\mathrm{T}_{L/F}(x)=x+\gamma(x)$ for $x\in L$. We let $\chi_{L/F}$ be the quadratic character associated to L/F, so that $\chi_{L/F}(x)=(x,D)_F$ for $x\in F^\times$.

2.2 2×2 symmetric matrices

Let $a,b,c\in F$ and set

$$S=\begin{bmatrix} a & b/2\\ b/2 & c\end{bmatrix}. \tag{11}$$

Let $D=b^2/4-ac=-\det(S)$. Assume that $D\neq 0$. The *discriminant* $\mathrm{disc}(S)$ of S is the class in $F^\times/F^{\times 2}$ determined by D. It is known that there exists $g\in\mathrm{GL}(2,F)$ such that ${}^tgSg$ is of the form $\left[\begin{smallmatrix} a_1 & \\ & a_2\end{smallmatrix}\right]$ and that $(a_1,a_2)_F$ is independent of the choice of g such that ${}^tgSg$ is diagonal; we define the *Hasse* *invariant* $\varepsilon(S) \in \{\pm 1\}$ by $\varepsilon(S) = (a_1, a_2)_F$. In fact, one has:

S	g	${}^t gSg$	$\mathrm{disc}(S)$	$\varepsilon(S)$
$a \neq 0, c \neq 0$	$\begin{bmatrix} 1 & \frac{-b}{2a} \\ & 1 \end{bmatrix}$	$\begin{bmatrix} a & \\ & c-\frac{b^2}{4a} \end{bmatrix}$	$(\frac{b^2}{4} - ac)F^{\times 2}$	$(a, \frac{b^2}{4} - ac)_F = (c, \frac{b^2}{4} - ac)_F$
$a \neq 0, c = 0$	$\begin{bmatrix} 1 & \frac{-b}{2a} \\ & 1 \end{bmatrix}$	$\begin{bmatrix} a & \\ & -\frac{b^2}{4a} \end{bmatrix}$	$F^{\times 2}$	1
$a = 0, c \neq 0$	$\begin{bmatrix} & 1 \\ 1 & -\frac{b}{2c} \end{bmatrix}$	$\begin{bmatrix} c & \\ & -\frac{b^2}{4c} \end{bmatrix}$	$F^{\times 2}$	1
$a = 0, c = 0$	$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$	$\begin{bmatrix} b & \\ & -b \end{bmatrix}$	$F^{\times 2}$	1

If $\mathrm{disc}(S) = F^{\times 2}$, then we say that S is *split*. If S is split, then for any $\lambda \in F^\times$ there exists $g \in \mathrm{GL}(2, F)$ such that ${}^t gSg = \begin{bmatrix} & \lambda \\ \lambda & \end{bmatrix}$.

2.3 Another F-algebra

Let S be as in (11) with $\mathrm{disc}(S) \neq 0$. Set $D = b^2/4 - ac$. We define

$$A = A_S = \{ \begin{bmatrix} x-yb/2 & -ya \\ yc & x+yb/2 \end{bmatrix} : x, y \in F \}. \tag{12}$$

Then, with respect to matrix addition and multiplication, A is a two-dimensional F-algebra naturally containing F. One can verify that

$$A = \begin{bmatrix} & 1 \\ 1 & \end{bmatrix} \{ g \in \mathrm{M}_2(F) : {}^t gSg = \det(g)S \} \begin{bmatrix} & 1 \\ 1 & \end{bmatrix}. \tag{13}$$

We define $T = T_S = A^\times$. Let L be the quadratic extension associated to D; we also say that L is the *quadratic extension associated to* S. We define an isomorphism of F-algebras,

$$A \xrightarrow{\sim} L, \qquad \begin{bmatrix} x-yb/2 & -ya \\ yc & x+yb/2 \end{bmatrix} \longmapsto x + y\Delta. \tag{14}$$

The restriction of this isomorphism to T is an isomorphism $T \xrightarrow{\sim} L^\times$, and we identify characters of T and characters of $L^\times$ via this isomorphism. The automorphism of A corresponding to the automorphism γ of L will also be denoted by γ. It has the effect of replacing y by $-y$ in the matrix (12). We have $\det(t) = \mathrm{N}_{L/F}(t)$ for $t \in A$, where we identify elements of A and L via (14).

2.3.1 LEMMA. *Let T be as above, and assume that L is a field. Let B_2 be the group of upper triangular matrices in* $\mathrm{GL}(2, F)$. *Then* $TB_2 = \mathrm{GL}(2, F)$.

Proof. This can easily be verified using the explicit form of the matrices in T and the assumption $D \notin F^{\times 2}$. ∎

2.4 Bessel functionals

Let a, b and c be in F. Define S as in (11), and define a character $\theta = \theta_{a,b,c} = \theta_S$ of N by

$$\theta(\begin{bmatrix} 1 & & y & z \\ & 1 & x & y \\ & & 1 & \\ & & & 1 \end{bmatrix}) = \psi(ax + by + cz) = \psi(\mathrm{tr}(S[{}_{1}{}^{1}][\begin{smallmatrix} y & z \\ x & y \end{smallmatrix}])) \tag{15}$$

for $x, y, z \in F$. Every character of N is of this form for uniquely determined a, b, c in F, or, alternatively, for a uniquely determined symmetric 2×2 matrix S. We say that θ is *non-degenerate* if $\det(S) \neq 0$. Given S with $\det(S) \neq 0$, let A be as in (12), and let $T = A^\times$. We embed T into $\mathrm{GSp}(4, F)$ via the map defined by

$$t \longmapsto \begin{bmatrix} t & \\ & \det(t)t' \end{bmatrix}, \qquad t \in T. \tag{16}$$

The image of T in $\mathrm{GSp}(4, F)$ will also be denoted by T; the usage should be clear from the context. For $t \in T$ we have $\lambda(t) = \det(t) = \mathrm{N}_{L/F}(t)$. It is easily verified that

$$\theta(tnt^{-1}) = \theta(n) \qquad \text{for } n \in N \text{ and } t \in T.$$

We refer to the semidirect product

$$D = TN \tag{17}$$

as the *Bessel subgroup* defined by character θ (or, the matrix S). Given a character Λ of T (identified with a character of $L^\times$ as explained above), we can define a character $\Lambda \otimes \theta$ of D by

$$(\Lambda \otimes \theta)(tn) = \Lambda(t)\theta(n) \qquad \text{for } n \in N \text{ and } t \in T.$$

Every character of D whose restriction to N coincides with θ is of this form for an appropriate Λ.

Now let (π, V) be an admissible representation of $\mathrm{GSp}(4, F)$. Let θ be a non-degenerate character of N, and let Λ be a character of the associated group T. We say that π admits a (Λ, θ)-*Bessel functional* if $\mathrm{Hom}_D(V, \mathbb{C}_{\Lambda\otimes\theta}) \neq 0$. A non-zero element β of $\mathrm{Hom}_D(V, \mathbb{C}_{\Lambda\otimes\theta})$ is called a (Λ, θ)-*Bessel functional* for π. If such a β exists, then π admits a model consisting of functions $B : \mathrm{GSp}(4, F) \to \mathbb{C}$ with the Bessel transformation property

$$B(tng) = \Lambda(t)\theta(n)B(g) \qquad \text{for } t \in T,\ n \in N \text{ and } g \in \mathrm{GSp}(4, F),$$

by associating to each v in V the function B_v that is defined by $B_v(g) = \beta(\pi(g)v)$ for $g \in \mathrm{GSp}(4, F)$. We note that if π admits a central character ω_π and a (Λ, θ)-Bessel functional, then $\Lambda|_{F^\times} = \omega_\pi$. For a character σ of $F^\times$, it is easy to verify that

$$\mathrm{Hom}_D(\pi, \mathbb{C}_{\Lambda\otimes\theta}) = \mathrm{Hom}_D(\sigma\pi, \mathbb{C}_{(\sigma\circ\mathrm{N}_{L/F})\Lambda\otimes\theta}). \tag{18}$$

If π is irreducible, then, using that $\pi^\vee \cong \omega_\pi^{-1}\pi$ (Proposition 2.3 of [37]), one can also verify that

$$\mathrm{Hom}_D(\pi, \mathbb{C}_{\Lambda\otimes\theta}) \cong \mathrm{Hom}_D(\pi^\vee, \mathbb{C}_{(\Lambda\circ\gamma)^{-1}\otimes\theta}). \tag{19}$$

The twisted Jacquet module of V with respect to N and θ is the quotient $V_{N,\theta} = V/V(N,\theta)$, where $V(N,\theta)$ is the subspace spanned by all vectors $\pi(n)v-\theta(n)v$ for v in V and n in N. This Jacquet module carries an action of T induced by the representation π. Evidently, there is a natural isomorphism

$$\mathrm{Hom}_D(V,\mathbb{C}_{\Lambda\otimes\theta}) \cong \mathrm{Hom}_T(V_{N,\theta},\mathbb{C}_\Lambda). \tag{20}$$

Hence, when calculating the possible Bessel functionals on a representation (π,V), a first step often consists in calculating the Jacquet modules $V_{N,\theta}$. We will use this method to calculate the possible Bessel functionals for most of the non-supercuspidal, irreducible, admissible representations of $\mathrm{GSp}(4,F)$. The few representations that are inaccessible with this method will be treated using the theta correspondence.

In this paper we do not assume that (Λ,θ)-Bessel functionals are unique up to scalars. See Sect. 6.3 for some remarks on uniqueness.

Finally, instead of $\mathrm{GSp}(4,F)$ as defined in this paper, in the literature it is common to work with the group G' of $g\in\mathrm{GL}(4,F)$ such that ${}^tg\left[\begin{smallmatrix}&1_2\\-1_2&\end{smallmatrix}\right]g = \lambda(g)\left[\begin{smallmatrix}&1_2\\-1_2&\end{smallmatrix}\right]$ for some $\lambda(g)\in F^\times$. For the convenience of the reader, we will explain how to translate statements about Bessel functionals from this paper into statements using G'. The groups $\mathrm{GSp}(4,F)$ and G' are isomorphic via the map $i:\mathrm{GSp}(4,F)\longrightarrow G'$ defined by $i(g)=LgL$ for $g\in\mathrm{GSp}(4,F)$, where $L=\left[\begin{smallmatrix}&1&&\\1&&&\\&&1&\\&&&1\end{smallmatrix}\right]$. We note that ${}^tL=L=L^{-1}$, $L^2=1$, and the inverse of i is given by $i^{-1}(g')=Lg'L$ for $g'\in G'$. If H is a subgroup of $\mathrm{GSp}(4,F)$, then we define $H'=i(H)$, and refer to H' as the subgroup of G' *corresponding* to H. For example, the subgroup N' of G' corresponding to N is

$$N'=\{\left[\begin{smallmatrix}1&&x&y\\&1&y&z\\&&1&\\&&&1\end{smallmatrix}\right]:x,y,z\in F\}.$$

If π is a smooth representation of a subgroup H of $\mathrm{GSp}(4,F)$ on a complex vector space V, then we define the representation π' of H' on V *corresponding* to π by the formula $\pi'(g')=\pi(i^{-1}(g'))$ for $g'\in H'$. Now let $S=\left[\begin{smallmatrix}a&b/2\\b/2&c\end{smallmatrix}\right]$ be as above, with $\det(S)\neq 0$. The character $\theta'=\theta'_S$ of N' corresponding to the character θ_S of N is given by the formula

$$\theta'(\left[\begin{smallmatrix}1&&x&y\\&1&y&z\\&&1&\\&&&1\end{smallmatrix}\right])=\psi(ax+by+cz)=\psi(\mathrm{tr}(S\left[\begin{smallmatrix}x&y\\y&z\end{smallmatrix}\right]))$$

for $x,y,z\in F$. The subgroup $T'=T'_S$ of G' corresponding to $T=T_S$ is

$$T'=\{\left[\begin{smallmatrix}t&\\&\det(t)\cdot{}^tt^{-1}\end{smallmatrix}\right]:t\in\mathrm{GL}(2,F):{}^ttSt=\det(t)S\}.$$

More explicitly, the group of $t\in\mathrm{GL}(2,F)$ such that ${}^ttSt=\det(t)S$ consists of the matrices

$$t=\left[\begin{smallmatrix}x+yb/2&yc\\-ya&x-yb/2\end{smallmatrix}\right] \tag{21}$$

where $x, y \in F$, $x^2-y^2(b^2/4-ac) \neq 0$, with $b^2/4-ac = -\det(S)$, as usual. With L as above, there is an isomorphism $T' \xrightarrow{\sim} L^\times$ given by $\left[\begin{smallmatrix} t & \\ & \det(t)\cdot {}^t t^{-1}\end{smallmatrix}\right] \mapsto x+y\Delta$ for t as in (21). Suppose that Λ is a character of $L^\times$; identify Λ with a character of T as explained above. The corresponding character Λ' of T' is given by the formula

$$\Lambda'(\left[\begin{smallmatrix} t & \\ & \det(t)\cdot {}^t t^{-1}\end{smallmatrix}\right]) = \Lambda(x+y\Delta)$$

for t as in (21). Finally, suppose that (π, V) is an admissible representation of $\mathrm{GSp}(4, F)$, and let π' be the representation of G' on V corresponding to π. There is an equality

$$\mathrm{Hom}_{D'}(V, \mathbb{C}_{\Lambda'\otimes\theta'}) = \mathrm{Hom}_D(V, \mathbb{C}_{\Lambda\otimes\theta}),$$

with $D' = T'N'$. The non-zero elements of $\mathrm{Hom}_{D'}(V, \mathbb{C}_{\Lambda'\otimes\theta'})$ are called (Λ', θ')-Bessel functionals for π', and the last equality asserts that the set of (Λ', θ')-Bessel functionals for π' is the same as the set of (Λ, θ)-Bessel functionals for π.

2.5 Action on Bessel functionals

There is an action of M, defined in (2), on the set of Bessel functionals. Let (π, V) be an irreducible, admissible representation of $\mathrm{GSp}(4, F)$, and let $\beta : V \to \mathbb{C}$ be a (Λ, θ)-Bessel functional for π. Let $a, b, c \in F$ be such that (15) holds. Let $m \in M$, with $m = \left[\begin{smallmatrix} g & \\ & \lambda g' \end{smallmatrix}\right]$, where $\lambda \in F^\times$ and $g \in \mathrm{GL}(2, F)$. Define $m \cdot \beta : V \to \mathbb{C}$ by $(m \cdot \beta)(v) = \beta(\pi(m^{-1})v)$ for $v \in V$. Calculations show that $m \cdot \beta$ is a (Λ', θ')-Bessel functional with θ' defined by

$$\theta'(\left[\begin{smallmatrix} 1 & & y & z \\ & 1 & x & y \\ & & 1 & \\ & & & 1 \end{smallmatrix}\right]) = \psi(a'x + b'y + c'z) = \psi(\mathrm{tr}(S'\left[\begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix}\right]\left[\begin{smallmatrix} y & z \\ x & y \end{smallmatrix}\right])), \quad x, y, z \in F,$$

where

$$S' = \left[\begin{smallmatrix} a' & b'/2 \\ b'/2 & c' \end{smallmatrix}\right] = \lambda\, {}^t h S h \quad \text{with} \quad h = \left[\begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix}\right] g^{-1} \left[\begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix}\right].$$

Since $\mathrm{disc}(S') = \mathrm{disc}(S)$, the quadratic extension L' associated to S' is the same as the quadratic extension L associated to S. There is an isomorphism of F-algebras

$$A' = A_{S'} \xrightarrow{\sim} A = A_S, \qquad a \mapsto g^{-1} a g.$$

Let $T' = A'^\times$. Finally, $\Lambda' : T' \to \mathbb{C}^\times$ is given by $\Lambda'(t') = \Lambda(g^{-1}t'g)$ for $t' \in T'$. For example, assume that β' is a *split* Bessel functional, i.e., a Bessel functional for which the discriminant of the associated symmetric matrix S' is the class $F^{\times 2}$. By Sect. 2.2 there exists m as above such that $\beta' = m \cdot \beta$, where the symmetric matrix S associated to the (Λ, θ)-Bessel functional β is

$$S = \left[\begin{smallmatrix} & 1/2 \\ 1/2 & \end{smallmatrix}\right], \tag{22}$$

and

$$\theta(\begin{bmatrix} 1 & & y & z \\ & 1 & x & y \\ & & 1 & \\ & & & 1 \end{bmatrix}) = \psi(y). \tag{23}$$

In this case

$$T = T_S = \{\begin{bmatrix} a & & & \\ & b & & \\ & & a & \\ & & & b \end{bmatrix} :\ a, b \in F^\times\}. \tag{24}$$

Sometimes when working with split Bessel functionals it is more convenient to work with the conjugate group

$$N_{\rm alt} = s_2^{-1} N s_2 = \begin{bmatrix} 1 & * & & * \\ & 1 & & \\ & * & 1 & * \\ & & & 1 \end{bmatrix} \tag{25}$$

and the conjugate character

$$\theta_{\rm alt}(\begin{bmatrix} 1 & -y & & z \\ & 1 & & \\ & x & 1 & y \\ & & & 1 \end{bmatrix}) = \psi(y). \tag{26}$$

In this case the stabilizer of $\theta_{\rm alt}$ is

$$T_{\rm alt} = \{\begin{bmatrix} a & & & \\ & a & & \\ & & b & \\ & & & b \end{bmatrix} :\ a, b \in F^\times\}. \tag{27}$$

2.6 Galois conjugation of Bessel functionals

The action of M can be used to define the Galois conjugate of a Bessel functional. Let S be as in (11), and let $A = A_S$ and $T = T_S$. Define

$$h_\gamma = \begin{cases} \begin{bmatrix} 1 & b/a \\ & -1 \end{bmatrix} & \text{if } a \neq 0, \\ \begin{bmatrix} 1 & \\ -b/c & -1 \end{bmatrix} & \text{if } a = 0 \text{ and } c \neq 0, \\ \begin{bmatrix} & 1 \\ 1 & \end{bmatrix} & \text{if } a = c = 0. \end{cases} \tag{28}$$

Then $h_\gamma \in {\rm GL}(2, F)$, $h_\gamma^2 = 1$, $S = {}^t h_\gamma S h_\gamma$ and $\det(h_\gamma) = -1$. Set

$$g_\gamma = \begin{bmatrix} & 1 \\ 1 & \end{bmatrix} h_\gamma^{-1} \begin{bmatrix} & 1 \\ 1 & \end{bmatrix} = \begin{bmatrix} & 1 \\ 1 & \end{bmatrix} h_\gamma \begin{bmatrix} & 1 \\ 1 & \end{bmatrix} \in {\rm GL}(2, F), \quad m_\gamma = \begin{bmatrix} g_\gamma & \\ & g'_\gamma \end{bmatrix} \in M.$$

We have $g_\gamma T g_\gamma^{-1} = T$, and the diagrams

$$\begin{tikzcd} A \arrow[r, "\sim"] \arrow[d, "\text{conjugation by } g_\gamma"'] & L \arrow[d, "\gamma"] \\ A \arrow[r, "\sim"] & L \end{tikzcd} \qquad \begin{tikzcd} T \arrow[r, "\sim"] \arrow[d, "\text{conjugation by } g_\gamma"'] & L^\times \arrow[d, "\gamma"] \\ T \arrow[r, "\sim"] & L^\times \end{tikzcd}$$

commute. Let (π, V) be an irreducible, admissible representation of ${\rm GSp}(4, F)$, and let β be a (Λ, θ)-Bessel functional for π. We refer to $m_\gamma \cdot \beta$ as the *Galois conjugate* of β. We note that $m_\gamma \cdot \beta$ is a $(\Lambda \circ \gamma, \theta)$-Bessel functional for π. Hence,

$${\rm Hom}_D(\pi, \mathbb{C}_{\Lambda \otimes \theta}) \cong {\rm Hom}_D(\pi, \mathbb{C}_{(\Lambda \circ \gamma) \otimes \theta}). \tag{29}$$

In combination with (19), we get

$$\mathrm{Hom}_D(\pi, \mathbb{C}_{\Lambda\otimes\theta}) \cong \mathrm{Hom}_D(\pi^\vee, \mathbb{C}_{\Lambda^{-1}\otimes\theta}). \tag{30}$$

2.7 Waldspurger functionals

Our analysis of Bessel functionals will often involve a similar type of functional on representations of $\mathrm{GL}(2,F)$. Let θ and S be as in (15), and let $T \cong L^\times$ be the associated subgroup of $\mathrm{GL}(2,F)$. Let Λ be a character of T. Let (π, V) be an irreducible, admissible representation of $\mathrm{GL}(2,F)$. A (Λ,θ)-*Waldspurger functional* on π is a non-zero linear map $\beta: V \to \mathbb{C}$ such that

$$\beta(\pi(g)v) = \Lambda(g)\beta(v) \qquad \text{for all } v \in V \text{ and } g \in T.$$

For trivial Λ, such functionals were the subject of Proposition 9 of [39] and Proposition 8 of [40]. For general Λ see [38], [34] and Lemme 8 of [40]. The (Λ,θ)-Waldspurger functionals are the non-zero elements of the space $\mathrm{Hom}_T(\pi, \mathbb{C}_\Lambda)$, and it is known that this space is at most one-dimensional. An obvious necessary condition for $\mathrm{Hom}_T(\pi, \mathbb{C}_\Lambda) \neq 0$ is that $\Lambda\big|_{F^\times}$ equals ω_π, the central character of π. By Sect. 2.6, Galois conjugation on T is given by conjugation by an element of $\mathrm{GL}(2,F)$. Hence,

$$\mathrm{Hom}_T(\pi, \mathbb{C}_\Lambda) \cong \mathrm{Hom}_T(\pi, \mathbb{C}_{\Lambda\circ\gamma}). \tag{31}$$

Using $\pi^\vee \cong \omega_\pi^{-1}\pi$, one verifies that

$$\mathrm{Hom}_T(\pi, \mathbb{C}_\Lambda) \cong \mathrm{Hom}_T(\pi^\vee, \mathbb{C}_{(\Lambda\circ\gamma)^{-1}}). \tag{32}$$

In combination with (31), we also have

$$\mathrm{Hom}_T(\pi, \mathbb{C}_\Lambda) \cong \mathrm{Hom}_T(\pi^\vee, \mathbb{C}_{\Lambda^{-1}}). \tag{33}$$

Let π^{JL} denote the Jacquet-Langlands lifting of π in the case that π is a discrete series representation, and 0 otherwise. Then, by the discussion on p. 1297 of [38],

$$\dim \mathrm{Hom}_T(\pi, \mathbb{C}_\Lambda) + \dim \mathrm{Hom}_T(\pi^{\mathrm{JL}}, \mathbb{C}_\Lambda) = 1. \tag{34}$$

It is easy to see that, for any character σ of $F^\times$,

$$\mathrm{Hom}_T(\pi, \mathbb{C}_\Lambda) = \mathrm{Hom}_T(\sigma\pi, \mathbb{C}_{(\sigma\circ \mathrm{N}_{L/F})\Lambda}). \tag{35}$$

For Λ such that $\Lambda\big|_{F^\times} = \sigma^2$, it is known that

$$\dim(\mathrm{Hom}_T(\sigma\mathrm{St}_{\mathrm{GL}(2)}, \mathbb{C}_\Lambda)) = \begin{cases} 0 & \text{if } L \text{ is a field and } \Lambda = \sigma\circ\mathrm{N}_{L/F}, \\ 1 & \text{otherwise;} \end{cases} \tag{36}$$

see Proposition 1.7 and Theorem 2.4 of [38]. As in the case of Bessel functionals, we call a Waldspurger functional *split* if the discriminant of the associated matrix S lies in $F^{\times 2}$. By Lemme 8 of [40], an irreducible, admissible, infinite-dimensional representation of $\mathrm{GL}(2,F)$ admits a split (Λ,θ)-Waldspurger functional with respect to any character Λ of T that satisfies $\Lambda\big|_{F^\times} = \omega_\pi$ (this can also be proved in a way analogous to the proof of Proposition 3.4.2 below, utilizing the standard zeta integrals for $\mathrm{GL}(2)$).

3 Split Bessel functionals

Irreducible, admissible, generic representations of $\mathrm{GSp}(4,F)$ admit a theory of zeta integrals, and every zeta integral gives rise to a split Bessel functional. As a consequence, generic representations admit *all* possible split Bessel functionals; see Proposition 3.4.2 below for a precise formulation.
To put the theory of zeta integrals on a solid foundation, we will use P_3-theory. The group P_3, defined below, plays a role in the representation theory of $\mathrm{GSp}(4)$ similar to the "mirabolic" subgroup in the theory for $\mathrm{GL}(n)$. Some of what follows is a generalization of Sects. 2.5 and 2.6 of [28], where P_3-theory was developed under the assumption of trivial central character. The general case requires only minimal modifications.
While every generic representation admits split Bessel functionals, we will see that the converse is not true. P_3-theory can also be used to identify the non-generic representations that admit a split Bessel functional. This is explained in Sect. 3.5 below.

3.1 The group P_3 and its representations

Let P_3 be the subgroup of $\mathrm{GL}(3,F)$ defined as the intersection

$$P_3 = \mathrm{GL}(3,F) \cap \begin{bmatrix} * & * & * \\ * & * & * \\ & & 1 \end{bmatrix}.$$

We recall some facts about this group, following [3]. Let

$$U_3 = P_3 \cap \begin{bmatrix} 1 & * & * \\ & 1 & * \\ & & 1 \end{bmatrix}, \qquad N_3 = P_3 \cap \begin{bmatrix} 1 & & * \\ & 1 & * \\ & & 1 \end{bmatrix}.$$

We define characters Θ and Θ' of U_3 by

$$\Theta(\begin{bmatrix} 1 & u_{12} & * \\ & 1 & u_{23} \\ & & 1 \end{bmatrix}) = \psi(u_{12}+u_{23}), \qquad \Theta'(\begin{bmatrix} 1 & u_{12} & * \\ & 1 & u_{23} \\ & & 1 \end{bmatrix}) = \psi(u_{23}).$$

If (π, V) is a smooth representation of P_3, we may consider the twisted Jacquet modules

$$V_{U_3,\Theta} = V/V(U_3,\Theta), \qquad V_{U_3,\Theta'} = V/V(U_3,\Theta')$$

where $V(U_3,\Theta)$ (resp. $V(U_3,\Theta')$) is spanned by all elements of the form $\pi(u)v - \Theta(u)v$ (resp. $\pi(u)v - \Theta'(u)v$) for v in V and u in U_3. Note that $V_{U_3,\Theta'}$ carries an action of the subgroup $\begin{bmatrix} * & & \\ & 1 & \\ & & 1 \end{bmatrix} \cong F^\times$ of P_3. We may also consider the Jacquet module $V_{N_3} = V/V(N_3)$, where $V(N_3)$ is the space spanned by all vectors of the form $\pi(u)v - v$ for v in V and u in N_3. Note that V_{N_3} carries an action of the subgroup $\begin{bmatrix} * & * & \\ * & * & \\ & & 1 \end{bmatrix} \cong \mathrm{GL}(2,F)$ of P_3.
Next we define three classes of smooth representations of P_3, associated with the groups $\mathrm{GL}(0)$, $\mathrm{GL}(1)$ and $\mathrm{GL}(2)$. Let

$$\tau_{\mathrm{GL}(0)}^{P_3}(1) := \text{c-Ind}_{U_3}^{P_3}(\Theta), \tag{37}$$

where c-Ind denotes compact induction. Then $\tau^{P_3}_{\mathrm{GL}(0)}(1)$ is a smooth, irreducible representation of P_3. Next, let χ be a smooth representation of $\mathrm{GL}(1,F) \cong F^\times$. Define a representation $\chi \otimes \Theta'$ of the subgroup $\left[\begin{smallmatrix} * & * & * \\ & 1 & * \\ & & 1 \end{smallmatrix}\right]$ of P_3 by

$$(\chi \otimes \Theta')(\left[\begin{smallmatrix} a & * & * \\ & 1 & y \\ & & 1 \end{smallmatrix}\right]) = \chi(a)\psi(y).$$

Then

$$\tau^{P_3}_{\mathrm{GL}(1)}(\chi) := \text{c-Ind}^{P_3}_{\left[\begin{smallmatrix} * & * & * \\ & 1 & * \\ & & 1 \end{smallmatrix}\right]}(\chi \otimes \Theta')$$

is a smooth representation of P_3. It is irreducible if and only if χ is one-dimensional. Finally, let ρ be a smooth representation of $\mathrm{GL}(2,F)$. We define the representation $\tau^{P_3}_{\mathrm{GL}(2)}(\rho)$ of P_3 to have the same space as ρ, and action given by

$$\tau^{P_3}_{\mathrm{GL}(2)}(\rho)(\left[\begin{smallmatrix} a & b & * \\ c & d & * \\ & & 1 \end{smallmatrix}\right]) = \rho(\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]). \tag{38}$$

Evidently, $\tau^{P_3}_{\mathrm{GL}(2)}(\rho)$ is irreducible if and only if ρ is irreducible.

3.1.1 Proposition. *Let notations be as above.*

i) Every irreducible, smooth representation of P_3 is isomorphic to exactly one of

$$\tau^{P_3}_{\mathrm{GL}(0)}(1), \qquad \tau^{P_3}_{\mathrm{GL}(1)}(\chi), \qquad \tau^{P_3}_{\mathrm{GL}(2)}(\rho),$$

where χ is a character of $F^\times$ and ρ is an irreducible, admissible representation of $\mathrm{GL}(2,F)$. Moreover, the equivalence classes of χ and ρ are uniquely determined.

ii) Let (π, V) be a smooth representation of P_3 of finite length. Then there exists a chain of P_3 subspaces

$$0 \subset V_2 \subset V_1 \subset V_0 = V$$

with the following properties,

$$\begin{aligned} V_2 &\cong \dim(V_{U_3,\Theta}) \cdot \tau^{P_3}_{\mathrm{GL}(0)}(1), \\ V_1/V_2 &\cong \tau^{P_3}_{\mathrm{GL}(1)}(V_{U_3,\Theta'}), \\ V_0/V_1 &\cong \tau^{P_3}_{\mathrm{GL}(2)}(V_{N_3}). \end{aligned}$$

Proof. See 5.1 – 5.15 of [3]. ∎

3.2 P_3-theory for arbitrary central character

It is easy to verify that any element of the Klingen parabolic subgroup Q can be written in a unique way as

$$\begin{bmatrix} ad-bc &&& \\ & a & b & \\ & c & d & \\ &&& 1 \end{bmatrix} \begin{bmatrix} 1 & -y & x & z \\ & 1 && x \\ && 1 & y \\ &&& 1 \end{bmatrix} \begin{bmatrix} u &&& \\ & u && \\ && u & \\ &&& u \end{bmatrix} \tag{39}$$

with $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right] \in \mathrm{GL}(2,F)$, $x,y,z \in F$, and $u \in F^\times$. Let Z^J be the center of the Jacobi group, consisting of all elements of $\mathrm{GSp}(4)$ of the form

$$\begin{bmatrix} 1 &&& * \\ & 1 && \\ && 1 & \\ &&& 1 \end{bmatrix}. \tag{40}$$

Evidently, Z^J is a normal subgroup of Q with $Z^J \cong F$. Let (π, V) be a smooth representation of $\mathrm{GSp}(4,F)$. Let $V(Z^J)$ be the span of all vectors $v - \pi(z)v$, where v runs through V and z runs through Z^J. Then $V(Z^J)$ is preserved by the action of Q. Hence Q acts on the quotient $V_{Z^J} := V/V(Z^J)$. Let $\bar{Q}$ be the subgroup of Q consisting of all elements of the form (39) with $u = 1$, i.e.,

$$\bar{Q} = \mathrm{GSp}(4) \cap \begin{bmatrix} * & * & * & * \\ & * & * & * \\ & * & * & * \\ &&& 1 \end{bmatrix}.$$

The map

$$i(\begin{bmatrix} ad-bc &&& \\ & a & b & \\ & c & d & \\ &&& 1 \end{bmatrix} \begin{bmatrix} 1 & -y & x & z \\ & 1 && x \\ && 1 & y \\ &&& 1 \end{bmatrix}) = \begin{bmatrix} a & b & \\ c & d & \\ && 1 \end{bmatrix} \begin{bmatrix} 1 && x \\ & 1 & y \\ && 1 \end{bmatrix} \tag{41}$$

establishes an isomorphism $\bar{Q}/Z^J \cong P_3$.

Recall the character ψ_{c_1,c_2} of U defined in (8). Note that U maps onto U_3 under the map (41), and that the diagrams

$$\begin{tikzcd} U \arrow[r, "i"] \arrow[dr, "\psi_{-1,1}"'] & U_3 \arrow[d, "\Theta"] \\ & \mathbb{C}^\times \end{tikzcd} \qquad\qquad \begin{tikzcd} U \arrow[r, "i"] \arrow[dr, "\psi_{-1,0}"'] & U_3 \arrow[d, "\Theta'"] \\ & \mathbb{C}^\times \end{tikzcd}$$

are commutative. The radical N_Q (see (6)) maps onto N_3 under the map (41). The following theorem is exactly like Theorem 2.5.3 of [28], except that the hypothesis of trivial central character is removed.

3.2.1 THEOREM. *Let (π, V) be an irreducible, admissible representation of $\mathrm{GSp}(4,F)$. The quotient $V_{Z^J} = V/V(Z^J)$ is a smooth representation of $\bar{Q}/Z^J$, and hence, via the map (41), defines a smooth representation of P_3. As a representation of P_3, V_{Z^J} has finite length. Hence, V_{Z^J} has a finite filtration by P_3 subspaces such that the successive quotients are irreducible and of the form $\tau_{\mathrm{GL}(0)}^{P_3}(1)$, $\tau_{\mathrm{GL}(1)}^{P_3}(\chi)$ or $\tau_{\mathrm{GL}(2)}^{P_3}(\rho)$ for some character χ of $F^\times$, or some irreducible, admissible representation ρ of $\mathrm{GL}(2,F)$. Moreover, the following statements hold:*

i) There exists a chain of P_3 *subspaces*

$$0 \subset V_2 \subset V_1 \subset V_0 = V_{Z^J}$$

such that

$$\begin{aligned} V_2 &\cong \dim \mathrm{Hom}_U(V, \psi_{-1,1}) \cdot \tau^{P_3}_{\mathrm{GL}(0)}(1), \\ V_1/V_2 &\cong \tau^{P_3}_{\mathrm{GL}(1)}(V_{U,\psi_{-1,0}}), \\ V_0/V_1 &\cong \tau^{P_3}_{\mathrm{GL}(2)}(V_{N_Q}). \end{aligned}$$

Here, the vector space $V_{U,\psi_{-1,0}}$ *admits a smooth action of* $\mathrm{GL}(1,F) \cong F^\times$ *induced by the operators*

$$\pi(\begin{bmatrix} a & & & \\ & a & & \\ & & 1 & \\ & & & 1 \end{bmatrix}), \quad a \in F^\times,$$

and V_{N_Q} *admits a smooth action of* $\mathrm{GL}(2,F)$ *induced by the operators*

$$\pi(\begin{bmatrix} \det g & & \\ & g & \\ & & 1 \end{bmatrix}), \quad g \in \mathrm{GL}(2,F).$$

ii) The representation π *is generic if and only if* $V_2 \neq 0$*, and if* π *is generic, then* $V_2 \cong \tau^{P_3}_{\mathrm{GL}(0)}(1)$*.*

iii) We have $V_2 = V_{Z^J}$ *if and only if* π *is supercuspidal. If* π *is supercuspidal and generic, then* $V_{Z^J} = V_2 \cong \tau^{P_3}_{\mathrm{GL}(0)}(1)$ *is non-zero and irreducible. If* π *is supercuspidal and non-generic, then* $V_{Z^J} = V_2 = 0$*.*

Proof. This is an application of Proposition 3.1.1. See Theorem 2.5.3 of [28] for the details of the proof. ∎

Given an irreducible, admissible representation (π, V) of $\mathrm{GSp}(4,F)$, one can calculate the semisimplifications of the quotients V_0/V_1 and V_1/V_2 in the P_3-filtration from the Jacquet modules of π with respect to the Siegel and Klingen parabolic subgroups. The results are exactly the same as in Appendix A.4 of [28] (where it was assumed that π has trivial central character).
Note that there is a typo in Table A.5 of [28]: The entry for Vd in the "s.s.(V_0/V_1)" column should be $\tau^{P_3}_{\mathrm{GL}(2)}(\nu(\nu^{-1/2}\sigma \times \nu^{-1/2}\xi\sigma))$.

3.3 Generic representations and zeta integrals

Let π be an irreducible, admissible, generic representation of $\mathrm{GSp}(4,F)$. Recall from Sect. 1.2 that $\mathcal{W}(\pi, \psi_{c_1,c_2})$ denotes the Whittaker model of π with respect to the character ψ_{c_1,c_2} of U. For W in $\mathcal{W}(\pi, \psi_{c_1,c_2})$ and $s \in \mathbb{C}$, we define the *zeta integral* $Z(s, W)$ by

$$Z(s,W) = \int\limits_{F^\times} \int\limits_{F} W(\begin{bmatrix} a & & & \\ & a & & \\ & x & 1 & \\ & & & 1 \end{bmatrix}) |a|^{s-3/2} \, dx \, d^\times a. \tag{42}$$

It was proved in Proposition 2.6.3 of [28] that there exists a real number s_0, independent of W, such that $Z(s,W)$ converges for $\Re(s) > s_0$ to an element of $\mathbb{C}(q^{-s})$. In particular, all zeta integrals have meromorphic continuation to all of $\mathbb{C}$. Let $I(\pi)$ be the $\mathbb{C}$-vector subspace of $\mathbb{C}(q^{-s})$ spanned by all $Z(s,W)$ for W in $\mathcal{W}(\pi,\psi_{c_1,c_2})$. It is easy to see that $I(\pi)$ is independent of the choice of ψ and c_1, c_2 in $F^\times$.

3.3.1 Proposition. *Let π be a generic, irreducible, admissible representation of $\mathrm{GSp}(4,F)$. Then $I(\pi)$ is a non-zero $\mathbb{C}[q^{-s},q^s]$-module containing $\mathbb{C}$, and there exists $R(X) \in \mathbb{C}[X]$ such that $R(q^{-s})I(\pi) \subset \mathbb{C}[q^{-s},q^s]$, so that $I(\pi)$ is a fractional ideal of the principal ideal domain $\mathbb{C}[q^{-s},q^s]$ whose quotient field is $\mathbb{C}(q^{-s})$. The fractional ideal $I(\pi)$ admits a generator of the form $1/Q(q^{-s})$ with $Q(0) = 1$, where $Q(X) \in \mathbb{C}[X]$.*

Proof. The proof is almost word for word the same as that of Proposition 2.6.4 of [28]. The only difference is that, in the calculation starting at the bottom of p. 79 of [28], the element q is taken from $\bar{Q}$ instead of Q. ∎

The quotient $1/Q(q^{-s})$ in this proposition is called the *L-factor* of π, and denoted by $L(s,\pi)$. If π is supercuspidal, then $L(s,\pi) = 1$. The L-factors for all irreducible, admissible, generic, non-supercuspidal representations are listed in Table A.8 of [28]. By definition,

$$\frac{Z(s,W)}{L(s,\pi)} \in \mathbb{C}[q^s,q^{-s}] \tag{43}$$

for all W in $\mathcal{W}(\pi,\psi_{c_1,c_2})$.

3.4 Generic representations admit split Bessel functionals

In this section we will prove that an irreducible, admissible, generic representation of $\mathrm{GSp}(4,F)$ admits split Bessel functionals with respect to *all* characters Λ of T. This is a characteristic feature of generic representations, which will follow from Proposition 3.5.1 in the next section.

3.4.1 Lemma. *Let (π,V) be an irreducible, admissible, generic representation of $\mathrm{GSp}(4,F)$. Let σ be a unitary character of $F^\times$, and let $s \in \mathbb{C}$ be arbitrary. Then there exists a non-zero functional $f_{s,\sigma} : V \to \mathbb{C}$ with the following properties.*

i) For all $x,y,z \in F$ and $v \in V$,

$$f_{s,\sigma}(\pi(\begin{bmatrix} 1 & & y & z \\ & 1 & x & y \\ & & 1 & \\ & & & 1 \end{bmatrix})v) = \psi(y) f_{s,\sigma}(v). \tag{44}$$

ii) For all $a \in F^\times$ and $v \in V$,

$$f_{s,\sigma}(\pi(\begin{bmatrix} a & & & \\ & 1 & & \\ & & a & \\ & & & 1 \end{bmatrix})v) = \sigma(a)^{-1}|a|^{-s+1/2} f_{s,\sigma}(v). \tag{45}$$

Proof. We may assume that $V = \mathcal{W}(\pi, \psi_{c_1,c_2})$ with $c_1 = 1$. Let $s_0 \in \mathbb{R}$ be such that $Z(s, W)$ is absolutely convergent for $\Re(s) > s_0$. Then the integral

$$Z_\sigma(s, W) = \int\limits_{F^\times}\int\limits_{F} W(\begin{bmatrix} a & & & \\ & a & & \\ & x & 1 & \\ & & & 1 \end{bmatrix})|a|^{s-3/2}\sigma(a)\,dx\,d^\times a \tag{46}$$

is also absolutely convergent for $\Re(s) > s_0$, since σ is unitary. Note that these are the zeta integrals for the twisted representation $\sigma\pi$. Therefore, by (43), the quotient $Z_\sigma(s, W)/L(s, \sigma\pi)$ is in $\mathbb{C}[q^{-s}, q^s]$ for all $W \in \mathcal{W}(\pi, \psi_{c_1,c_2})$. We may therefore define, for any complex s,

$$f_{s,\sigma}(W) = \frac{Z_\sigma(s, \pi(s_2)W)}{L(s, \sigma\pi)}, \tag{47}$$

where s_2 is as in (7). Straightforward calculations using the definition (46) show that (44) and (45) are satisfied for $\Re(s) > s_0$. Since both sides depend holomorphically on s, these identities hold on all of $\mathbb{C}$. ∎

3.4.2 Proposition. *Let (π, V) be an irreducible, admissible and generic representation of $\mathrm{GSp}(4, F)$. Let ω_π be the central character of π. Then π admits a split (Λ, θ)-Bessel functional with respect to any character Λ of T that satisfies $\Lambda\big|_{F^\times} = \omega_\pi$.*

Proof. Let θ be as in (23) with T as in (24). Let $s \in \mathbb{C}$ and σ be a unitary character of $F^\times$ such that

$$\Lambda(\begin{bmatrix} a & & & \\ & 1 & & \\ & & a & \\ & & & 1 \end{bmatrix}) = \sigma(a)^{-1}|a|^{-s+1/2} \qquad \text{for all } a \in F^\times.$$

Let $f_{s,\sigma}$ be as in Lemma 3.4.1. By (45),

$$f_{s,\sigma}(\pi(\begin{bmatrix} a & & & \\ & 1 & & \\ & & a & \\ & & & 1 \end{bmatrix})v) = \Lambda(a) f_{s,\sigma}(v) \qquad \text{for all } a \in F^\times. \tag{48}$$

Since $\Lambda\big|_{F^\times} = \omega_\pi$ we have in fact $f_{s,\sigma}(\pi(t)v) = \Lambda(t) f_{s,\sigma}(v)$ for all $t \in T$. Hence $f_{s,\sigma}$ is a Bessel functional as desired. ∎

3.5 Split Bessel functionals for non-generic representations

The converse of Proposition 3.4.2 is not true: There exist irreducible, admissible, non-generic representations of $\mathrm{GSp}(4,F)$ which admit split Bessel functionals. This follows from the following proposition. In fact, using this result and the P_3-filtrations listed in Table A.6 of [28], one can precisely identify which non-generic representations admit split Bessel functionals. Other than in the generic case, the possible characters Λ of T are restricted to a finite number.

3.5.1 PROPOSITION. *Let (π, V) be an irreducible, admissible and non-generic representation of $\mathrm{GSp}(4,F)$. Let the semisimplification of the quotient $V_1 = V_1/V_2$ in the P_3-filtration of π be given by $\sum_{i=1}^n \tau_{\mathrm{GL}(1)}^{P_3}(\chi_i)$ with characters χ_i of $F^\times$.*

i) π admits a split Bessel functional if and only if the quotient V_1 in the P_3-filtration of π is non-zero.

ii) Let β be a non-zero (Λ,θ)-Bessel functional, with θ as in (23), and a character Λ of the group T explicitly given in (24). Then there exists an i for which

$$\Lambda(\begin{bmatrix} a & & & \\ & 1 & & \\ & & a & \\ & & & 1 \end{bmatrix}) = |a|^{-1}\chi_i(a) \qquad \text{for all} \quad a \in F^\times. \tag{49}$$

iii) If V_1 is non-zero, then there exists an i such that π admits a split (Λ,θ)-Bessel functional with respect to a character Λ of T satisfying (49).

iv) The space of split (Λ,θ)-Bessel functionals is zero or one-dimensional.

v) The representation π does not admit any split Bessel functionals if and only if π is of type IVd, Vd, VIb, VIIIb, IXb, or is supercuspidal.

Proof. Let N_{alt} be as in (25) and θ_{alt} be as in (26). We use the fact that any $(\Lambda,\theta_{\mathrm{alt}})$-Bessel functional factors through the twisted Jacquet module $V_{N_{\mathrm{alt}},\theta_{\mathrm{alt}}}$. To calculate this Jacquet module, we use the P_3-filtration of Theorem 3.2.1. Since π is non-generic, the P_3-filtration simplifies to

$$0 \subset V_1 \subset V_0 = V_{Z^J},$$

with V_1 of type $\tau_{\mathrm{GL}(1)}^{P_3}$ and V_0/V_1 of type $\tau_{\mathrm{GL}(2)}^{P_3}$. Taking further twisted Jacquet modules and observing Lemma 2.5.6 of [28], it follows that

$$V_{N_{\mathrm{alt}},\theta_{\mathrm{alt}}} = (V_1)_{\left[\begin{smallmatrix} 1 & & \\ * & 1 & * \\ & & 1 \end{smallmatrix}\right],\psi}, \qquad \text{where} \quad \psi(\left[\begin{smallmatrix} 1 & & \\ x & 1 & y \\ & & 1 \end{smallmatrix}\right]) = \psi(y).$$

By Lemma 2.5.5 of [28], after suitable renaming,

$$0 = J_n \subset \ldots \subset J_1 \subset J_0 = (V_1)_{\left[\begin{smallmatrix} 1 & & \\ * & 1 & * \\ & & 1 \end{smallmatrix}\right],\psi},$$

where J_i/J_{i+1} is one-dimensional, and $\mathrm{diag}(a,1,1)$ acts on J_i/J_{i+1} by $|a|^{-1}\chi_i(a)$. Table A.6 of [28] shows that all the χ_i are pairwise distinct. This proves i), ii), iii) and iv).
v) If π is one of the representations mentioned in v), then $V_1/V_2 = 0$ by Theorem 3.2.1 (in the supercuspidal case), or by Table A.6 in [28] (in the non-supercuspidal case). By part i), π does not admit a split Bessel functional. For any representation not mentioned in v), the quotient V_1/V_2 is non-zero, so that a split Bessel functional exists by iii). ∎

4 Theta correspondences

Let S be as in (11), and let $\theta = \theta_S$ be as in (15). Let (π, V) be an irreducible, admissible representation of $\mathrm{GSp}(4,F)$, and let (σ, W) be an irreducible, admissible representation of $\mathrm{GO}(X)$, where X is an even-dimensional, symmetric, bilinear space. Let ω be the Weil representation of the group R, consisting of the pairs $(g,h) \in \mathrm{GSp}(4,F) \times \mathrm{GO}(X)$ with the same similitude factors, on the Schwartz space $\mathcal{S}(X^2)$. Assume that the pair (π,σ) occurs in the theta correspondence defined by ω, i.e., $\mathrm{Hom}_R(\omega, \pi\otimes\sigma) \neq 0$. It is a theme in the theory of the theta correspondence to relate the twisted Jacquet module $V_{N,\theta}$ of π to invariant functionals on σ; a necessary condition for the non-vanishing of $V_{N,\theta}$ is that X represents S. See for example the remarks in Sect. 6 of [26].
Applications to (Λ, θ_S)-Bessel functionals also require the involvement of T. The idea is roughly as follows. The group T is contained in M. Moreover, $\omega(m,h)$ for (m,h) in $R \cap (M \times \mathrm{GO}(X))$ is given by an action of such pairs on X^2. The study of this action leads to the definition of certain compatible embeddings of T into $\mathrm{GO}(X)$. Using these embeddings allows us to show that if π has a (Λ,θ)-Bessel functional, then σ admits a non-zero functional transforming according to Λ^{-1}.
After setting up notations and studying the embeddings of T mentioned above, we obtain the main result of this section, Theorem 4.4.6. Section 4.7 contains the applications to Bessel functionals.

4.1 The spaces

In this section we will consider non-degenerate symmetric bilinear spaces $(X, \langle\cdot,\cdot\rangle)$ over F such that

$$\dim X = 2, \text{ or } \dim X = 4 \text{ and } \mathrm{disc}(X) = 1. \tag{50}$$

We begin by recalling the constructions of the isomorphism classes of these spaces, and the characterization of their similitude groups. Let $m \in F^\times$, $A = A_{\left[\begin{smallmatrix}1&\\&-m\end{smallmatrix}\right]}$ and $T = A^\times$ be as in Sect. 2.3, so that

$$A = \{\left[\begin{smallmatrix}x&-y\\-ym&x\end{smallmatrix}\right] : x,y \in F\}, \quad T = A^\times = \{\left[\begin{smallmatrix}x&-y\\-ym&x\end{smallmatrix}\right] : x,y\in F,\ x^2-y^2m \neq 0\}. \tag{51}$$

Let $\lambda \in F^\times$. Define a non-degenerate two-dimensional symmetric bilinear space $(X_{m,\lambda}, \langle\cdot,\cdot\rangle_{m,\lambda})$ by

$$X_{m,\lambda} = A_{\left[\begin{smallmatrix}1&\\&-m\end{smallmatrix}\right]}, \qquad \langle x_1, x_2\rangle_{m,\lambda} = \lambda\,\mathrm{tr}(x_1 x_2^*)/2, \quad x_1, x_2 \in X_{m,\lambda}. \tag{52}$$

Here, $*$ is the canonical involution of 2×2 matrices, given by $\left[\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right]^* = \left[\begin{smallmatrix}d&-b\\-c&a\end{smallmatrix}\right]$. Define a homomorphism $\rho : T \to \mathrm{GSO}(X_{m,\lambda})$ by $\rho(t)x = tx$ for $x \in X_{m,\lambda}$. We also recall the Galois conjugation map $\gamma : A \to A$ from Sect. 2.3; it is given by $\gamma(x) = x^*$ for $x \in A$. The map γ can be regarded as an F linear endomorphism

$$\gamma : X_{m,\lambda} \longrightarrow X_{m,\lambda}, \tag{53}$$

and as such is contained in $\mathrm{O}(X_{m,\lambda})$ but not in $\mathrm{SO}(X_{m,\lambda})$.

4.1.1 Lemma. *If $(X_{m,\lambda}, \langle\cdot,\cdot\rangle_{m,\lambda})$ is as in (52), then $\mathrm{disc}(X_{m,\lambda}) = mF^{\times 2}$, $\varepsilon(X_{m,\lambda}) = (\lambda, m)$, and the homomorphism ρ is an isomorphism, so that*

$$\rho : T \xrightarrow{\sim} \mathrm{GSO}(X_{m,\lambda}). \tag{54}$$

The image $\rho(T)$ and the map γ generate $\mathrm{GO}(X_{m,\lambda})$. If $(X_{m,\lambda}, \langle\cdot,\cdot\rangle_{m,\lambda})$ and $(X_{m',\lambda'}, \langle\cdot,\cdot\rangle_{m',\lambda'})$ are as in (52), then $(X_{m,\lambda}, \langle\cdot,\cdot\rangle_{m,\lambda}) \cong (X_{m',\lambda'}, \langle\cdot,\cdot\rangle_{m',\lambda'})$ if and only if $mF^{\times 2} = m'F^{\times 2}$ and $(\lambda, m)_F = (\lambda', m')_F$. Every two-dimensional, non-degenerate symmetric bilinear space over F is isomorphic $(X_{m,\lambda}, \langle\cdot,\cdot\rangle_{m,\lambda})$ for some m and λ.

Proof. Let $m, \lambda \in F^\times$. In $X_{m,\lambda}$ let $x_1 = \left[\begin{smallmatrix}1&\\&1\end{smallmatrix}\right]$ and $x_2 = \left[\begin{smallmatrix}&1\\m&\end{smallmatrix}\right]$. Then x_1, x_2 is a basis for $X_{m,\lambda}$, and in this basis the matrix for $X_{m,\lambda}$ is $\lambda\left[\begin{smallmatrix}1&\\&-m\end{smallmatrix}\right]$. Calculations using this matrix show that $\mathrm{disc}(X_{m,\lambda}) = mF^{\times 2}$ and $\varepsilon(X_{m,\lambda}) = (\lambda, m)_F$. The map ρ is clearly injective. To see that ρ is surjective, let $h \in \mathrm{GSO}(X_{m,\lambda})$. Write h in the ordered basis x_1, x_2 so that $h = \left[\begin{smallmatrix}h_1&h_2\\h_3&h_4\end{smallmatrix}\right]$. By the definition of $\mathrm{GSO}(X_{m,\lambda})$, we have ${}^t h\lambda\left[\begin{smallmatrix}1&\\&-m\end{smallmatrix}\right]h = \det(h)\lambda\left[\begin{smallmatrix}1&\\&-m\end{smallmatrix}\right]$. By the definition of T, this implies that $t = \left[\begin{smallmatrix}&1\\1&\end{smallmatrix}\right]h\left[\begin{smallmatrix}&1\\1&\end{smallmatrix}\right] \in T$. Hence, $h = \left[\begin{smallmatrix}h_1&h_3m\\h_3&h_1\end{smallmatrix}\right]$ for some $h_1, h_3 \in F$. Calculations now show that $\rho(t)x_1 = h(x_1)$ and $\rho(t)x_2 = h(x_2)$, so that $\rho(t) = h$. This proves the first assertion. The second assertion follows from the fact that two non-degenerate symmetric bilinear spaces over F with the same finite dimension are isomorphic if and only if they have the same discriminant and Hasse invariant. For the final assertion, let $(X, \langle\cdot,\cdot\rangle)$ be a two-dimensional, non-degenerate symmetric bilinear space over F. There exists a basis for X with respect to which the matrix for X is of the form $\left[\begin{smallmatrix}\alpha_1&\\&\alpha_2\end{smallmatrix}\right]$ for some $\alpha_1, \alpha_2 \in F^\times$. Then $\mathrm{disc}(X) = -\alpha_1\alpha_2 F^{\times 2}$ and $\varepsilon(X) = (\alpha_1, \alpha_2)_F$. An argument shows that there exists $\lambda \in F^\times$ such that $(\lambda, \mathrm{disc}(X))_F = \varepsilon(X)$. We now have $(X, \langle\cdot,\cdot\rangle) \cong (X_{m,\lambda}, \langle\cdot,\cdot\rangle_{m,\lambda})$ with $m = \mathrm{disc}(X)$ because both spaces have the same discriminant and Hasse invariant. ∎

Next, define a four-dimensional non-degenerate symmetric bilinear space over F by setting

$$X_{\mathrm{M}_2} = \mathrm{M}_2(F), \qquad \langle x_1, x_2\rangle_{\mathrm{M}_2} = \mathrm{tr}(x_1 x_2^*)/2, \quad x_1, x_2 \in X_{\mathrm{M}_2}. \tag{55}$$

Here, $*$ is the canonical involution of 2×2 matrices, given by $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]^* = \left[\begin{smallmatrix} d & -b \\ -c & a \end{smallmatrix}\right]$. Define $\rho : \mathrm{GL}(2,F)\times\mathrm{GL}(2,F) \to \mathrm{GSO}(X_{\mathrm{M}_2})$ by $\rho(g_1,g_2)x = g_1 x g_2^*$ for $g_1, g_2 \in \mathrm{GL}(2,F)$ and $x \in X_{\mathrm{M}_2}$. The map $* : X_{\mathrm{M}_2} \to X_{\mathrm{M}_2}$ is contained in $\mathrm{O}(X_{\mathrm{M}_2})$ but not in $\mathrm{SO}(X_{\mathrm{M}_2})$.

Finally, let H be the division quaternion algebra over F. Let $1, i, j, k$ be a quaternion algebra basis for H, i.e.,

$$H = F + Fi + Fj + Fk, \quad i^2 \in F^\times,\ j^2 \in F^\times,\ k = ij,\ ij = -ji. \tag{56}$$

Let $*$ be the canonical involution on H so that $(a + b\cdot i + c\cdot j + d\cdot k)^* = a - b\cdot i - c\cdot j - d\cdot k$, and define the norm and trace functions $\mathrm{N}, \mathrm{T} : H \to F$ by $\mathrm{N}(x) = xx^*$ and $\mathrm{T}(x) = x + x^*$ for $x \in H$. Define another four-dimensional non-degenerate symmetric bilinear space over F by setting

$$X_H = H, \qquad \langle x_1, x_2\rangle_H = \mathrm{T}(x_1 x_2^*)/2, \quad x_1, x_2 \in X_H. \tag{57}$$

Define $\rho : H^\times \times H^\times \to \mathrm{GSO}(X_H)$ by $\rho(h_1,h_2)x = h_1 x h_2^*$ for $h_1, h_2 \in H^\times$ and $x \in X_H$. The map $* : X_H \to X_H$ is contained in $\mathrm{O}(X_H)$ but not in $\mathrm{SO}(X_H)$.

4.1.2 Lemma. *The symmetric bilinear space $(X_{\mathrm{M}_2}, \langle\cdot,\cdot\rangle_{\mathrm{M}_2})$ is non-degenerate, has dimension four, discriminant $\mathrm{disc}(X_{\mathrm{M}_2}) = 1$, and Hasse invariant $\varepsilon(X_{\mathrm{M}_2}) = (-1,-1)_F$. The $(X_H, \langle\cdot,\cdot\rangle_H)$ symmetric bilinear space is non-degenerate, has dimension four, discriminant $\mathrm{disc}(X_H) = 1$, and Hasse invariant $\varepsilon(X_H) = -(-1,-1)_F$. The sequences*

$$1 \longrightarrow F^\times \longrightarrow \mathrm{GL}(2,F)\times\mathrm{GL}(2,F) \xrightarrow{\rho} \mathrm{GSO}(X_{\mathrm{M}_2}) \longrightarrow 1, \tag{58}$$

$$1 \longrightarrow F^\times \longrightarrow H^\times \times H^\times \xrightarrow{\rho} \mathrm{GSO}(X_H) \longrightarrow 1 \tag{59}$$

are exact; here, the second maps send a to (a, a^{-1}) for $a \in F^\times$. The image $\rho(\mathrm{GL}(2,F)\times\mathrm{GL}(2,F))$ and the map $*$ generate $\mathrm{GO}(X_{\mathrm{M}_2})$, and the image $\rho(H^\times \times H^\times)$ and the map $*$ generate $\mathrm{GO}(X_H)$. Every four-dimensional, non-degenerate symmetric linear space over F of discriminant 1 is isomorphic to $(X_{\mathrm{M}_2}, \langle\cdot,\cdot\rangle_{\mathrm{M}_2})$ or $(X_H, \langle\cdot,\cdot\rangle_H)$.

Proof. See, for example, Sect. 2 of [27]. ∎

4.2 Embeddings

Suppose that $(X, \langle\cdot,\cdot\rangle)$ satisfies (50). We define an action of the group $\mathrm{GL}(2,F)\times\mathrm{GO}(X)$ on the set X^2 by

$$(g,h)\cdot(x_1,x_2) = (hx_1, hx_2)g^{-1} = (g_1' h x_1 + g_3' h x_2, g_2' h x_1 + g_4' h x_2) \tag{60}$$

for $(x_1,x_2) \in X^2$, $h \in \mathrm{GO}(X)$ and $g \in \mathrm{GL}(2,F)$ with $g^{-1} = \left[\begin{smallmatrix} g_1' & g_2' \\ g_3' & g_4' \end{smallmatrix}\right]$. For S as in (11) with $\det(S) \neq 0$, we define

$$\Omega = \Omega_S = \Omega_{S,(X,\langle\cdot,\cdot\rangle)} = \{(x_1,x_2) \in X^2 : \left[\begin{smallmatrix} \langle x_1,x_1\rangle & \langle x_1,x_2\rangle \\ \langle x_1,x_2\rangle & \langle x_2,x_2\rangle \end{smallmatrix}\right] = S\}. \tag{61}$$

We say that $(X, \langle\cdot,\cdot\rangle)$ *represents* S if the set Ω is non-empty.

4.2.1 Lemma. *Let $(X, \langle\cdot,\cdot\rangle)$ be a non-degenerate symmetric bilinear space over F satisfying (50), and let S be as in (11) with $\det(S) \neq 0$. The subgroup*

$$B = B_S = \{(g,h) \in \mathrm{GL}(2,F) \times \mathrm{GO}(X) : {}^tgSg = \det(g)S \text{ and } \det(g) = \lambda(h)\} \tag{62}$$

maps $\Omega = \Omega_S$ to itself under the action of $\mathrm{GL}(2,F) \times \mathrm{GO}(X)$ on X^2.

Proof. Let $(g,h) \in B$, and let $g = [\begin{smallmatrix} g_1 & g_2 \\ g_3 & g_4 \end{smallmatrix}]$. To start, we note that the assumption ${}^tgSg = \det(g)S$ is equivalent to ${}^tg^{-1}Sg^{-1} = \det(g)^{-1}S$, which is in turn equivalent to

$$\begin{aligned} ag_4^2 - bg_3g_4 + cg_3^2 &= \det(g)a, \\ -ag_4g_2 + b(g_1g_4 + g_2g_3)/2 - cg_3g_1 &= \det(g)b/2, \\ ag_2^2 - bg_2g_1 + cg_1^2 &= \det(g)c. \end{aligned}$$

Let $(x_1, x_2) \in \Omega$ and set $(y_1, y_2) = (g, h_1(t)) \cdot (x_1, x_2)$. By the definition of the action and Ω, and using $\det(g) = \lambda(h)$, we have

$$\begin{aligned} \langle y_1, y_1\rangle &= \det(g)^{-1}\big(g_4^2\langle x_1, x_1\rangle - 2g_3g_4\langle x_1, x_2\rangle + g_3^2\langle x_2, x_2\rangle\big) \\ &= \det(g)^{-1}\big(g_4^2a - g_3g_4b + g_3^2c\big) \\ &= a. \end{aligned}$$

Similarly, $\langle y_1, y_2\rangle = b/2$ and $\langle y_2, y_2\rangle = c$. It follows that $(y_1, y_2) = (g, h_1(t))(x_1, x_2) \in \Omega$. ∎

Let $(X, \langle\cdot,\cdot\rangle)$ be a non-degenerate symmetric bilinear space over F satisfying (50), and let S be as in (11) with $\det(S) \neq 0$. Assume that Ω is non-empty, and let $T = T_S$, as in Sect. 2.3. The goal of this section is to define, for each $z \in \Omega$, a set

$$\mathcal{E}(z) = \mathcal{E}_{(X,\langle\cdot,\cdot\rangle),S}(z) \tag{63}$$

of embeddings $\tau : T \to \mathrm{GSO}(X)$ such that:

$$\tau(t) = t \text{ for } t \in F^\times \subset T; \tag{64}$$

$$\lambda(\tau(t)) = \det(t) \text{ for } t \in T, \text{ so that } (t, \tau(t)) \in B \text{ for } t \in T; \tag{65}$$

$$([\begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix}]\, t\, [\begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix}], \tau(t)) \cdot z = z \text{ for } t \in T. \tag{66}$$

We begin by noting some properties of Ω. The set Ω is closed in X^2. The subgroup $\mathrm{O}(X) \cong 1 \times \mathrm{O}(X) \subset B \subset \mathrm{GL}(2,F) \times \mathrm{GO}(X)$ preserves Ω, i.e., if $h \in \mathrm{O}(X)$ and $(x_1, x_2) \in \Omega$, then $(hx_1, hx_2) \in \Omega$. Since $\det(S) \neq 0$, the group $\mathrm{O}(X)$ acts transitively on Ω. If $\dim X = 4$, then $\mathrm{SO}(X)$ acts transitively on Ω. If $\dim X = 2$, then the action of $\mathrm{SO}(X)$ on Ω has two orbits.

4.2.2 Lemma. *Let $(X, \langle\cdot,\cdot\rangle)$ be a non-degenerate symmetric bilinear space over F satisfying (50), and let S be as in (11) with $\det(S) \neq 0$. Assume that $\dim X = 2$ and Ω is non-empty. Let $z = (z_1, z_2) \in \Omega$. For*

$$t = \left[\begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix}\right] g \left[\begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix}\right] = \left[\begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix}\right]\left[\begin{smallmatrix} g_1 & g_2 \\ g_3 & g_4 \end{smallmatrix}\right]\left[\begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix}\right] \in T$$

let $\tau_z(t): X \to X$ be the linear map that has g as matrix in the ordered basis z_1, z_2 for X, so that

$$\tau_z(t)(z_1) = g_1 z_1 + g_3 z_2,$$
$$\tau_z(t)(z_2) = g_2 z_1 + g_4 z_2.$$

i) For $t \in T$, the map $\tau_z(t)$ is contained in $\mathrm{GSO}(X)$ and $\lambda(\tau_z(t)) = \det(t)$.

ii) If z' lies in the $\mathrm{SO}(X)$ orbit of z, and $t \in T$, then $\tau_z(t) = \tau_{z'}(t)$.

iii) The map sending t to $\tau_z(t)$ defines an isomorphism $\tau_z : T \xrightarrow{\sim} \mathrm{GSO}(X)$.

iv) Let $h_0 \in \mathrm{O}(X)$ with $\det(h_0) = -1$. Let $z' \in \Omega$ not be in the $\mathrm{SO}(X)$ orbit of z. Then $\tau_{z'}(t) = h_0 \tau_z(t) h_0^{-1}$ for $t \in T$.

v) Let $t \in T$. The element $(\left[\begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix}\right] t \left[\begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix}\right], \tau_z(t)) \in B$ acts by the identity on the $\mathrm{SO}(X)$ orbit of z, and maps the other $\mathrm{SO}(X)$ orbit of Ω to itself.

Proof. i) A computation verifies that $\tau_z(t) \in \mathrm{GO}(X)$, with similitude factor $\lambda(\tau_z(t)) = \det(g) = \det(t)$, and the equality $\det(\tau_z(t)) = \lambda(\tau_z(t))$ implies that $\tau_z(t) \in \mathrm{GSO}(X)$ by the definition of $\mathrm{GSO}(X)$.
ii) Suppose that $z' = (z_1', z_2')$ lies in the $\mathrm{SO}(X)$ orbit of z, and let $c \in \mathrm{SO}(X)$ be such that $c(z_1) = z_1'$ and $c(z_2) = z_2'$. Then $\tau_{z'}(t) = c\tau_z(t)c^{-1}$. But the group $\mathrm{GSO}(X)$ is abelian, so that $\tau_{z'}(t) = c\tau_z(t)c^{-1} = \tau_z(t)$.
iii) Calculations prove that $\tau_z : T \to \mathrm{GSO}(X)$ is an isomorphism.
iv) Let $z'' = h_0(z)$. A calculation shows that $\tau_{z''}(t) = h_0\tau_z(t)h_0^{-1}$ for $t \in T$. By, ii), $\tau_{z''}(t) = \tau_{z'}(t)$ for $t \in T$.
v) Write $g = \left[\begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix}\right] t \left[\begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix}\right]$, so that ${}^t gSg = \det(g)S$. Let $g = \left[\begin{smallmatrix} g_1 & g_2 \\ g_3 & g_4 \end{smallmatrix}\right]$. By the definition of $\tau_z(t)$, we have

$$\begin{aligned}(g, \tau_z(t)) \cdot z = (&\det(g)^{-1} g_4(g_1 z_1 + g_3 z_2) - \det(g)^{-1} g_3(g_2 z_1 + g_4 z_2),\\ &\det(g)^{-1}(-g_2)(g_1 z_1 + g_3 z_2) + \det(g)^{-1} g_1(g_2 z_1 + g_4 z_2))\\ = z&.\end{aligned}$$

By ii), it follows that $(g, \tau_z(t))$ acts by the identity on all of the $\mathrm{SO}(X)$ orbit of z. Next, let $z' \in \Omega$ with $z' \notin \mathrm{SO}(X)z$. Assume that $(g, \tau_z(t)) \cdot z' \in \mathrm{SO}(X)z$; we will obtain a contradiction. Since $(g, \tau_z(t)) \cdot z' \in \mathrm{SO}(X)z$ and since we have already proved that $(g, \tau_z(t))$ acts by the identity on $\mathrm{SO}(X)z$, we have:

$$(g, \tau_z(t)) \cdot \big((g, \tau_z(t)) \cdot z'\big) = (g, \tau_z(t)) \cdot z'$$
$$(g, \tau_z(t)) \cdot z' = z'.$$

This is a contradiction since $z' \notin \mathrm{SO}(X)z$ and $(g, \tau_z(t)) \cdot z' \in \mathrm{SO}(X)z$. ∎

Let $(X, \langle\cdot,\cdot\rangle)$ be a non-degenerate symmetric bilinear space over F satisfying (50), and let S be as in (11) with $\det(S) \neq 0$. Assume that $\dim X = 2$ and Ω is non-empty. For $z \in \Omega$, we define

$$\mathcal{E}(z) = \mathcal{E}_{(X,\langle\cdot,\cdot\rangle),S}(z) = \{\tau_z\}, \tag{67}$$

with τ_z as defined in Lemma 4.2.2. It is evident from Lemma 4.2.2 that the element of $\mathcal{E}(z)$ has the properties (64), (65), and (66).

4.2.3 LEMMA. *Let $(X, \langle\cdot,\cdot\rangle)$ be a non-degenerate symmetric bilinear space over F satisfying (50), and let S be as in (11) with $\det(S) \neq 0$. Assume that $\dim X = 2$. Let $\lambda, \lambda' \in F^\times$, and set $\Omega = \Omega_{\lambda S}$ and $\Omega' = \Omega_{\lambda' S}$. Assume that Ω and Ω' are non-empty. Then*

$$\bigcup_{z\in\Omega} \mathcal{E}(z) = \bigcup_{z'\in\Omega'} \mathcal{E}(z'). \tag{68}$$

Proof. Let Ω_1 and Ω_2 be the two $\mathrm{SO}(X)$ orbits of the action of $\mathrm{SO}(X)$ on Ω so that $\Omega = \Omega_1 \sqcup \Omega_2$, and analogously define and write $\Omega' = \Omega'_1 \sqcup \Omega'_2$. Let $z = (z_1, z_2) \in \Omega_1$ and $z' = (z'_1, z'_2) \in \Omega'_1$. Define a linear map $h : X \to X$ by setting $h(z_1) = z'_1$ and $h(z_2) = z'_2$. We have $\langle h(x), h(y)\rangle = (\lambda'/\lambda)\langle x, y\rangle$ for $x, y \in X$, so that $h \in \mathrm{GO}(X)$. Assume that $h \notin \mathrm{GSO}(X)$. Let $z'' = (z''_1, z''_2) \in \Omega'_2$, and let $h' : X \to X$ be the linear map defined by $h'(z'_1) = z''_1$ and $h'(z'_2) = z''_2$. Then $h' \in \mathrm{O}(X)$ with $\det(h') = -1$, so that $h'h \in \mathrm{GSO}(X)$ and $(h'h)(z_1) = z''_1$ and $(h'h)(z_2) = z''_2$. Therefore, by renumbering if necessary, we may assume that $h \in \mathrm{GSO}(X)$. Next, a calculation shows that $h\tau_z(t)h^{-1} = \tau_{z'}(t)$ for $t \in T$. Since $\mathrm{GSO}(X)$ is abelian, this means that $\tau_z = \tau_{z'}$. The claim (68) follows now from ii) and iv) of Lemma 4.2.2. ∎

4.2.4 LEMMA. *Let $(X, \langle\cdot,\cdot\rangle)$ be a non-degenerate symmetric bilinear space over F satisfying (50), and let S be as in (11) with $\det(S) \neq 0$. Assume that $\dim X = 4$ and Ω is non-empty. Let $z = (z_1, z_2) \in \Omega$, and set $U = Fz_1 + Fz_2$, so that $X = U \oplus U^\perp$ with $\dim U = \dim U^\perp = 2$. There exists $\lambda \in F^\times$ such that $(U^\perp, \langle\cdot,\cdot\rangle)$ represents λS.*

Proof. Let $\mathrm{M}_{4,1}(F)$ be the F vector space of 4×1 matrices with entries from F. Let $D = -\det(S)$. Let $\lambda \in F^\times$, and define a four-dimensional symmetric bilinear space X_λ by letting $X_\lambda = \mathrm{M}_{4,1}(F)$ with symmetric bilinear form b given by $b(x, y) = {}^t x M y$, where

$$M = \begin{bmatrix} S & \\ & \lambda S \end{bmatrix}.$$

Evidently, $\mathrm{disc}(X_\lambda) = 1$, and the Hasse invariant of X_λ is $\varepsilon(X_\lambda) = (-1,-1)_F(-\lambda, D)_F$. Now assume that X is isotropic. Then the Hasse invariant of X is $(-1,-1)_F$. It follows that if $\lambda = -1$, then $\varepsilon(X_\lambda) = \varepsilon(X)$, so that $X_\lambda \cong X$. By the Witt cancellation theorem, $(U^\perp, \langle\cdot,\cdot\rangle)$ represents λS. Next, assume that X is anisotropic, so that $\varepsilon(X) = -(-1,-1)_F$. By hypothesis, $(X, \langle\cdot,\cdot\rangle)$ represents S; since X is anisotropic, this implies that $D \notin F^{\times 2}$. Since, $D \notin F^{\times 2}$, there exists $\lambda \in F^\times$ such that $-1 = (-\lambda, D)_F$. It follows that $\varepsilon(X_\lambda) = \varepsilon(X)$, so that $X_\lambda \cong X$; again the Witt cancellation theorem implies that $(U^\perp, \langle\cdot,\cdot\rangle)$ represents λS. ∎

Let $(X, \langle\cdot,\cdot\rangle)$ be a non-degenerate symmetric bilinear space over F satisfying (50), and let S be as in (11) with $\det(S) \neq 0$. Assume that $\dim X = 4$, $\Omega = \Omega_S$ is non-empty, and let $T = T_S$, as in Sect. 2.3. Let $z = (z_1, z_2) \in \Omega$, and as in Lemma 4.2.4, let $U = Fz_1 + Fz_2$, so that $X = U \oplus U^\perp$ with $\dim U = \dim U^\perp = 2$. By Lemma 4.2.4 there exists $\lambda \in F^\times$ such that $(U^\perp, \langle\cdot,\cdot\rangle)$ represents λS. Let $\tau_z : T \to \mathrm{GSO}(U)$ be the isomorphism from Lemma 4.2.2 that is associated to z. Also, let $\tau_{z'}, \tau_{z''} : T \to \mathrm{GSO}(U^\perp)$ be the isomorphisms from Lemma 4.2.2, where z' and z'' are representatives for the two $\mathrm{SO}(U^\perp)$ orbits of $\mathrm{SO}(U^\perp)$ acting on $\Omega_{\lambda S,(U^\perp,\langle\cdot,\cdot\rangle)}$; by Lemma 4.2.3, $\{\tau_{z'}, \tau_{z''}\}$ does not depend on the choice of λ. We now define

$$\mathcal{E}(z) = \mathcal{E}_{(X,\langle\cdot,\cdot\rangle),S}(z) = \{\tau_1, \tau_2\}, \tag{69}$$

where $\tau_1, \tau_2 : T \to \mathrm{GSO}(X)$ are defined by

$$\tau_1(t) = \begin{bmatrix} \tau_z(t) & \\ & \tau_{z'}(t) \end{bmatrix}, \qquad \tau_2(t) = \begin{bmatrix} \tau_z(t) & \\ & \tau_{z''}(t) \end{bmatrix}$$

with respect to the decomposition $Z = U \oplus U^\perp$, for $t \in T$. For $t \in T$, the similitude factor of $\tau_i(t)$ is $\det(t)$. It is evident that the elements of $\mathcal{E}(z)$ satisfy (64), (65), and (66).

4.2.5 Lemma. *Let $(X, \langle\cdot,\cdot\rangle)$ be a non-degenerate symmetric bilinear space over F satisfying (50), and let S be as in (11) with $\det(S) \neq 0$. Assume that $\Omega = \Omega_S$ is non-empty, and let $A = A_S$ and $T = T_S$, as in Sect. 2.3. If $\dim X = 4$, assume that A is a field. Let $z \in \Omega$ and $\tau \in \mathcal{E}(z)$. Let C be a compact, open subset of Ω containing z. There exists a compact, open subset C_0 of Ω such that $z \in C_0 \subset C$ and $([\begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix}]\, t\, [\begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix}], \tau(t)) \cdot C_0 = C_0$ for $t \in T$.*

Proof. Assume $\dim X = 2$. Let C_0 be the intersection of C with the $\mathrm{SO}(X)$ orbit of z in Ω. Then C_0 is a compact, open subset of Ω because the $\mathrm{SO}(X)$ orbit of z in Ω is closed and open in Ω, and C is compact and open. We have $([\begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix}]\, t\, [\begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix}], \tau(t)) \cdot C_0 = C_0$ for $t \in T$ by v) of Lemma 4.2.2. Assume $\dim X = 4$. The group of pairs $([\begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix}]\, t\, [\begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix}], \tau(t))$ for $t \in T$ acts on X^2 and can be regarded as a subgroup of $\mathrm{GL}(X^2)$. The group T contains $F^\times$, and the pairs with $t \in F^\times$ act by the identity on X^2. The assumption that A is a field implies that $T/F^\times$ is compact, and hence the image $\mathcal{K}$ in $\mathrm{GL}(X^2)$ of this group of pairs is compact. There exists a lattice $\mathcal{L}$ of X^2 such that $k \cdot \mathcal{L} = \mathcal{L}$ for $k \in \mathcal{K}$. Also, by (66) we have that $k \cdot z = z$ for $k \in \mathcal{K}$. Let n be sufficiently large so that $(z + \varpi^n \mathcal{L}) \cap \Omega \subset C$. Then $C_0 = (z + \varpi^n \mathcal{L}) \cap \Omega$ is the desired set. ∎

4.3 Example embeddings

In this section we provide explicit formulas for the embeddings of the previous section.

4.3.1 Lemma. *Let S be as in (11). Let $m, \lambda \in F^\times$, and define $(X_{m,\lambda}, \langle\cdot,\cdot\rangle_{m,\lambda})$ as in (52). The set $\Omega = \Omega_S$ is non-empty if and only if $\mathrm{disc}(S) = mF^{\times 2}$ and $\varepsilon(S) = (\lambda, m)_F$. Assume that the set Ω is non-empty. Set $D = b^2/4 - ac$ so that $\mathrm{disc}(S) = DF^{\times 2}$, and define Δ and the quadratic extension $L = F + F\Delta$ of F (which need not be a field) associated to D as in Sect. 2.1. Similarly, define Δ_m with respect to m; the quadratic extension associated m is also L and $L = F + F\Delta_m$. The set of compositions*

$$L^\times \xrightarrow{\sim} T_S \xrightarrow{\tau} \mathrm{GSO}(X_{m,\lambda})$$

for $z \in \Omega$ and $\tau \in \mathcal{E}(z)$ is the same as the set consisting of the two compositions

$$L^\times \xrightarrow{\sim} T_{\left[\begin{smallmatrix}1&\\&-m\end{smallmatrix}\right]} \xrightarrow{\sim} \mathrm{GSO}(X_{m,\lambda}), \quad L^\times \xrightarrow{\gamma} L^\times \xrightarrow{\sim} T_{\left[\begin{smallmatrix}1&\\&-m\end{smallmatrix}\right]} \xrightarrow{\sim} \mathrm{GSO}(X_{m,\lambda}).$$

Here, the maps $L^\times \to T_{\left[\begin{smallmatrix}1&\\&-m\end{smallmatrix}\right]}$ are as in (14), and the isomorphism ρ of $T_{\left[\begin{smallmatrix}1&\\&-m\end{smallmatrix}\right]}$ with $\mathrm{GSO}(X_{m,\lambda})$ is as in (54).

Proof. By definition, Ω is non-empty if and only if there exist $x_1, x_2 \in X_{m,\lambda}$ such that $S = \left[\begin{smallmatrix}\langle x_1,x_1\rangle & \langle x_1,x_2\rangle\\ \langle x_1,x_2\rangle & \langle x_2,x_2\rangle\end{smallmatrix}\right]$. Since $X_{m,\lambda}$ is two-dimensional, this means that Ω is non-empty if and only if $(X_{m,\lambda}, \langle\cdot,\cdot\rangle_{m,\lambda})$ is equivalent to the symmetric bilinear space over F defined by S. From Lemma 4.1.1, we have $\mathrm{disc}(X_{m,\lambda}) = mF^{\times 2}$ and $\varepsilon(X_{m,\lambda}) = (\lambda, m)_F$. Since a finite-dimensional non-degenerate symmetric bilinear space over F is determined by its dimension, discriminant and Hasse invariant, it follows that Ω is non-empty if and only if $\mathrm{disc}(S) = mF^{\times 2}$ and $\varepsilon(S) = (\lambda, m)_F$.

Assume that Ω is non-empty, so that $\mathrm{disc}(S) = mF^{\times 2}$ and $\varepsilon(S) = (\lambda, m)_F$. Let $e \in F^\times$ be such that $\Delta = e\Delta_m$; then $b^2/4 - ac = D = e^2 m$. Assume first that $a \neq 0$. By Sect. 2.2, $\varepsilon(S) = (a, m)_F$. Therefore, $(a, m)_F = (\lambda, m)_F$. It follows that there exist $g, h \in F^\times$ such that $g^2 - mh^2 = \lambda^{-1}a$. Set

$$z_1 = \left[\begin{smallmatrix} g & h \\ -h(-m) & g \end{smallmatrix}\right], \quad z_2 = a^{-1}\left[\begin{smallmatrix} ehm+gb/2 & eg+hb/2 \\ -(eg+hb/2)(-m) & ehm+gb/2 \end{smallmatrix}\right].$$

Then $z_1, z_2 \in X_{m,\lambda}$, and a calculation shows that

$$\left[\begin{smallmatrix} \langle z_1,z_1\rangle_{m,\lambda} & \langle z_1,z_2\rangle_{m,\lambda} \\ \langle z_1,z_2\rangle_{m,\lambda} & \langle z_2,z_2\rangle_{m,\lambda} \end{smallmatrix}\right] = S.$$

It follows that $z = (z_1, z_2) \in \Omega$. Let $u \in L^\times$. Write $u = x + y\Delta$ for some $x, y \in F^\times$. By (14), u corresponds to $t = \left[\begin{smallmatrix} x-yb/2 & -ya \\ yc & x+yb/2 \end{smallmatrix}\right] \in T_S$. Using the definition of $\tau_z(t)$, we find that

$$\tau_z(t)(z_1) = \left[\begin{smallmatrix} gx-ehym & hx-egy \\ -(hx-egy)(-m) & gx-ehym \end{smallmatrix}\right], \tag{70}$$

$$\tau_z(t)(z_2) = \frac{1}{2a}\left[\begin{smallmatrix} (2ehm+bg)x-(behm+2e^2mg)y & (2eg+bh)x-(beg+2e^2mh)y \\ -((2eg+bh)x-(beg+2e^2mh)y)(-m) & (2ehm+bg)x-(behm+2e^2mg)y \end{smallmatrix}\right]. \tag{71}$$

On the other hand, we also have that $u = x + ye\Delta_m$, and u corresponds to the element $t' = \left[\begin{smallmatrix} x & -ye \\ (ye)(-m) & x \end{smallmatrix}\right]$ in $T_{\left[\begin{smallmatrix} 1 & \\ & -m \end{smallmatrix}\right]}$. Moreover, calculations show that $\rho(t')(z_1) = t' \cdot z_1$ and $\rho(t')(z_2) = t' \cdot z_2$ are as in (70) and (71), respectively, proving that the two compositions

$$L^\times \xrightarrow{\sim} T_S \xrightarrow{\tau_z} \mathrm{GSO}(X_{m,\lambda}), \qquad L^\times \xrightarrow{\sim} T_{\left[\begin{smallmatrix} 1 & \\ & -m \end{smallmatrix}\right]} \xrightarrow{\rho} \mathrm{GSO}(X_{m,\lambda})$$

are the same map. Next, let $z' = (\gamma(z_1), \gamma(z_2))$. Then $z' \in \Omega$, and calculations as above show that the two compositions

$$L^\times \xrightarrow{\sim} T_S \xrightarrow{\tau_{z'}} \mathrm{GSO}(X_{m,\lambda}), \qquad L^\times \xrightarrow{\gamma} L^\times \xrightarrow{\sim} T_{\left[\begin{smallmatrix} 1 & \\ & -m \end{smallmatrix}\right]} \xrightarrow{\rho} \mathrm{GSO}(X_{m,\lambda})$$

are the same. This completes the proof in this case since z and z' are representatives for the two $\mathrm{SO}(X_{m,\lambda})$ orbits of Ω, and by ii) of Lemma 4.2.2, $\cup_{w\in\Omega}\mathcal{E}(w) = \{\tau_z, \tau_{z'}\}$. Now assume that $a = 0$. Set

$$z_1 = \lambda^{-1}\left[\begin{smallmatrix} b/2 & -e \\ e(-m) & b/2 \end{smallmatrix}\right], \qquad z_2 = \left[\begin{smallmatrix} (c\lambda^{-1}+1)/2 & -eb^{-1}(c\lambda^{-1}-1) \\ eb^{-1}(c\lambda^{-1}-1)(-m) & (c\lambda^{-1}+1)/2 \end{smallmatrix}\right].$$

Again, a calculation shows that $z = (z_1, z_2) \in \Omega$. Let $u \in L^\times$ with $u = x + y\Delta$ for some $x, y \in F^\times$. Then u corresponds to $t = \left[\begin{smallmatrix} x-yb/2 & \\ yc & x+yb/2 \end{smallmatrix}\right] \in T_S$, and u corresponds to $t' = \left[\begin{smallmatrix} x & -ye \\ ye(-m) & x \end{smallmatrix}\right] \in T_{\left[\begin{smallmatrix} 1 & \\ & -m \end{smallmatrix}\right]}$. Computations show that $\tau_z(t)(z_1) = \rho(t')(z_1)$ and $\tau_z(t)(z_2) = \rho(t')(z_2)$, proving that the compositions

$$L^\times \xrightarrow{\sim} T_S \xrightarrow{\tau_{z'}} \mathrm{GSO}(X_{m,\lambda}), \qquad L^\times \xrightarrow{\sim} T_{\left[\begin{smallmatrix} 1 & \\ & -m \end{smallmatrix}\right]} \xrightarrow{\rho} \mathrm{GSO}(X_{m,\lambda})$$

are the same. As in the previous case, if $z' = (\gamma(z_1), \gamma(z_2))$, then $z' \in \Omega$, and the two compositions

$$L^\times \xrightarrow{\sim} T_S \xrightarrow{\tau_{z'}} \mathrm{GSO}(X_{m,\lambda}), \qquad L^\times \xrightarrow{\gamma} L^\times \xrightarrow{\sim} T_{\left[\begin{smallmatrix} 1 & \\ & -m \end{smallmatrix}\right]} \xrightarrow{\rho} \mathrm{GSO}(X_{m,\lambda})$$

are the same. As above, this completes the proof. ∎

Let $c \in F^\times$, and set

$$S = \left[\begin{smallmatrix} 1 & \\ & c \end{smallmatrix}\right]. \tag{72}$$

Let $(X_{\mathrm{M}_2}, \langle\cdot,\cdot\rangle_{\mathrm{M}_2})$ be as in (55). Let $A = A_S$ and $T = T_S$ be as in Sect. 2.3. We embed A in $\mathrm{M}_2(F)$ via the inclusion map. Set

$$z_1 = \left[\begin{smallmatrix} 1 & \\ & 1 \end{smallmatrix}\right], \quad z_2 = \left[\begin{smallmatrix} & 1 \\ -c & \end{smallmatrix}\right], \quad z_1' = \left[\begin{smallmatrix} 1 & \\ & -1 \end{smallmatrix}\right], \quad z_2' = \left[\begin{smallmatrix} & 1 \\ c & \end{smallmatrix}\right].$$

The vectors z_1, z_2, z_1', z_2' form an orthogonal ordered basis for X_{M_2}, and in this basis the matrix for X_{M_2} is

$$\begin{bmatrix} S & \\ & -S \end{bmatrix}.$$

As in Lemma 4.2.4, set $U = Fz_1 + Fz_2$. Then $U^\perp = Fz_1' + Fz_2'$, and the λ of Lemma 4.2.4 is -1. Calculations show that the set $\mathcal{E}(z) = \mathcal{E}_{X_{\mathrm{M}_2}}(z)$ of (69) is

$$\mathcal{E}_{X_{\mathrm{M}_2}}(z) = \{\tau_1, \tau_2\}, \qquad \tau_1(t) = \rho(t,1), \quad \tau_2(t) = \rho(1, \gamma(t)), \quad t \in T^\times. \tag{73}$$

Finally, let S be as in (72) with $-c \notin F^{\times 2}$, and let $(X_H, \langle \cdot, \cdot \rangle_H)$ be as in (57). Let $A = A_S$ and $T = T_S$ be as in Sect. 2.3. Let L be the quadratic extension associated to $-c$ as in Sect. 2.1; L is a field. Let e be a representative for the non-trivial coset of $F^\times / \mathrm{N}_{L/F}(L^\times)$, so that $(e, -c)_F = -1$. We realize the division quaternion algebra H over F as

$$H = F + Fi + Fj + Fk, \quad i^2 = -c, \; j^2 = e, \; k = ij, \; ij = -ji. \tag{74}$$

We embed A into H via the map defined by

$$\begin{bmatrix} x & -y \\ cy & x \end{bmatrix} \mapsto x - yi$$

for $x, y \in F$. Let

$$z_1 = 1, \quad z_2 = i, \quad z_1' = j, \quad z_2' = k. \tag{75}$$

The vectors z_1, z_2, z_1', z_2' form an orthogonal ordered basis for X_H, and in this basis the matrix for X_H is

$$\begin{bmatrix} S & \\ & -eS \end{bmatrix}.$$

As in Lemma 4.2.4, set $U = Fz_1 + Fz_2$. Then $U^\perp = Fz_1' + Fz_2'$, and the λ of Lemma 4.2.4 is $-e$. Calculations again show that the set $\mathcal{E}(z) = \mathcal{E}_{X_H}(z)$ of (69) is

$$\mathcal{E}_{X_H}(z) = \{\tau_1, \tau_2\}, \qquad \tau_1(t) = \rho(t,1), \quad \tau_2(t) = \rho(1, \gamma(t)), \quad t \in T^\times. \tag{76}$$

To close this subsection, we note that $(X_H, \langle \cdot, \cdot \rangle_H)$ does not represent S if S is as in (72) but $-c \in F^{\times 2}$. To see this, assume that $-c \in F^{\times 2}$ and $(X_H, \langle \cdot, \cdot \rangle_H)$ represents S; we will obtain a contradiction. Write $-c = t^2$ for some $t \in F^\times$. Since X_H represents S, there exist $x_1, x_2 \in H$ such that $\langle x_1, x_1 \rangle_H = \mathrm{N}(x_1) = 1$, $\langle x_2, x_2 \rangle_H = \mathrm{N}(x_2) = c = -t^2$ and $\langle x_1, x_2 \rangle_H = \mathrm{T}(x_1 x_2^*)/2 = 0$. A calculation shows that $\mathrm{N}(tx_1 + x_2) = 0$. Since H is a division algebra, this means that $tx_1 = -x_2$. Hence, $t^2 = N(tx_1) = \langle tx_1, tx_1 \rangle_H = \langle tx_1, -x_2 \rangle_H = -t \langle x_1, x_2 \rangle_H = 0$, a contradiction.

4.4 Theta correspondences and Bessel functionals

In this section we make the connection between Bessel functionals for $\mathrm{GSp}(4, F)$ and equivariant functionals on representations of $\mathrm{GO}(X)$. The main result is Theorem 4.4.6 below.

Let $(X, \langle\cdot,\cdot\rangle)$ be a non-degenerate symmetric bilinear space over F satisfying (50). We define the subgroup $\mathrm{GSp}(4,F)^+$ of $\mathrm{GSp}(4,F)$ by

$$\mathrm{GSp}(4,F)^+ = \{g \in \mathrm{GSp}(4,F) : \lambda(g) \in \lambda(\mathrm{GO}(X))\}. \tag{77}$$

The following lemma follows from (54) and the exact sequences (58) and (59), which facilitate the computation of $\lambda(\mathrm{GSO}(X))$. Note that $\mathrm{N}(H^\times) = F^\times$.

4.4.1 Lemma. *Let $(X, \langle\cdot,\cdot\rangle)$ be a non-degenerate symmetric bilinear space over F satisfying (50). Then*

$$[\mathrm{GSp}(4,F) : \mathrm{GSp}(4,F)^+] = \begin{cases} 1 & \text{if } \dim X = 4, \text{ or } \dim X = 2 \text{ and } \mathrm{disc}(X) = 1, \\ 2 & \text{if } \dim X = 2 \text{ and } \mathrm{disc}(X) \neq 1. \end{cases} \tag{78}$$

4.4.2 Lemma. *Let $(X, \langle\cdot,\cdot\rangle)$ be a non-degenerate symmetric bilinear space over F satisfying (50), and let S be as in (11) with $\det(S) \neq 0$. Let $\Omega = \Omega_S$ be as in (61), and assume that Ω is non-empty. Let $T = T_S$ be as in Sect. 2.3. Embed T as a subgroup of $\mathrm{GSp}(4,F)$, as in (16). Then T is contained in $\mathrm{GSp}(4,F)^+$.*

Proof. By (78) we may assume that $\dim X = 2$ and $\mathrm{disc}(X) \neq 1$. Since Ω is non-empty and $\dim X = 2$, we make take S to be the matrix of the symmetric bilinear form $\langle\cdot,\cdot\rangle$ on X. By definition, $\mathrm{GO}(X)$ is then the set of $h \in \mathrm{GL}(2,F)$ such that ${}^t hSh = \lambda(h)S$ for some $\lambda(h) \in F^\times$. From (13), we have that ${}^t hSh = \det(h)S$ for $h = \left[\begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix}\right] t \left[\begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix}\right]$ with $t \in T$. It follows that $\det(T)$ is contained in $\lambda(\mathrm{GO}(X))$. This implies that T is contained in $\mathrm{GSp}(4,F)^+$. ∎

Let $(X, \langle\cdot,\cdot\rangle)$ be a non-degenerate symmetric bilinear space over F satisfying (50). Define

$$R = \{(g,h) \in \mathrm{GSp}(4,F) \times \mathrm{GO}(X) : \lambda(g) = \lambda(h)\}.$$

We consider the Weil representation ω of R on the space $\mathcal{S}(X^2)$ defined with respect to ψ^2, where $\psi^2(x) = \psi(2x)$ for $x \in F$. If $\varphi \in \mathcal{S}(X^2)$, $g \in \mathrm{GL}(2,F)$ and $h \in \mathrm{GO}(X)$ with $\det(g) = \lambda(h)$, and $x_1, x_2 \in X$, then

$$\big(\omega(\begin{bmatrix} 1 & & y & z \\ & 1 & x & y \\ & & 1 & \\ & & & 1 \end{bmatrix}, 1)\varphi\big)(x_1,x_2) = \psi(\langle x_1,x_1\rangle x + 2\langle x_1,x_2\rangle y + \langle x_2,x_2\rangle z)\varphi(x_1,x_2), \tag{79}$$

$$\big(\omega(\left[\begin{smallmatrix} g & \\ & \det(g)g' \end{smallmatrix}\right], h)\varphi\big)(x_1,x_2) = (\det(g), \mathrm{disc}(X))_F\,\varphi((\left[\begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix}\right] g \left[\begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix}\right], h)^{-1} \cdot (x_1,x_2)). \tag{80}$$

For these formulas, see Sect. 1 of [27]; note that the additive character we are using is ψ^2. Also, in (80) we are using the action of $\mathrm{GL}(2,F) \times \mathrm{GO}(X)$ defined in (60).

We will also use the Weil representation ω_1 of

$$R_1 = \{(g,h) \in \mathrm{GL}(2,F) \times \mathrm{GO}(X) : \det(g) = \lambda(h)\}$$

on $\mathcal{S}(X)$ defined with respect to ψ^2. For formulas, again see Sect. 1 of [27]. The two Weil representations ω and ω_1 are related as follows.

4.4.3 Lemma. *The map*

$$T : \mathcal{S}(X) \otimes \mathcal{S}(X) \longrightarrow \mathcal{S}(X^2), \tag{81}$$

determined by the formula

$$T(\varphi_1 \otimes \varphi_2)(x_1, x_2) = \varphi_1(x_1)\varphi_2(x_2) \tag{82}$$

for φ_1 *and* φ_2 *in* $\mathcal{S}(X)$ *and* x_1 *and* x_2 *in* X, *is a well-defined complex linear isomorphism such that*

$$\begin{aligned} T \circ (\omega_1\big(\left[\begin{smallmatrix} a_2 & b_2 \\ c_2 & d_2 \end{smallmatrix}\right], h) \otimes \omega_1(\left[\begin{smallmatrix} a_1 & b_1 \\ c_1 & d_1 \end{smallmatrix}\right], h)\big) \\ = \omega(\begin{bmatrix} a_1 & & & b_1 \\ & a_2 & b_2 & \\ & c_2 & d_2 & \\ c_1 & & & d_1 \end{bmatrix}, h) \circ T \end{aligned} \tag{83}$$

for $g_1 = \left[\begin{smallmatrix} a_1 & b_1 \\ c_1 & d_1 \end{smallmatrix}\right]$ *and* $g_2 = \left[\begin{smallmatrix} a_2 & b_2 \\ c_2 & d_2 \end{smallmatrix}\right]$ *in* $\mathrm{GL}(2, F)$ *and* h *in* $\mathrm{GO}(X)$ *such that*

$$\det(g_1) = \det(g_2) = \lambda(h).$$

This lemma can be verified by a direct calculation using standard generators for $\mathrm{SL}(2, F)$.
Let $\theta = \theta_S$ be the character of N defined in (15) with respect to a matrix S as in (11). Let $\mathcal{S}(X^2)(N, \theta)$ be the subspace of $\mathcal{S}(X^2)$ spanned by all vectors $\omega(n)\varphi - \theta(n)\varphi$, where n runs through N and φ runs through $\mathcal{S}(X^2)$, and set $\mathcal{S}(X^2)_{N,\theta} = \mathcal{S}(X^2)/\mathcal{S}(X^2)(N, \theta)$.

4.4.4 Lemma. *(Rallis) Let* $(X, \langle\cdot,\cdot\rangle)$ *be a non-degenerate symmetric bilinear space over* F *satisfying* (50), *and let* S *be as in* (11) *with* $\det(S) \neq 0$. *If* $(X, \langle\cdot,\cdot\rangle)$ *does not represent* S, *then the twisted Jacquet module* $\mathcal{S}(X^2)_{N,\theta}$ *is zero. Assume that* $(X, \langle\cdot,\cdot\rangle)$ *represents* S. *The map* $\mathcal{S}(X^2) \to \mathcal{S}(\Omega)$ *defined by* $\varphi \mapsto \varphi|_\Omega$ *induces an isomorphism*

$$\mathcal{S}(X^2)_{N,\theta} \xrightarrow{\sim} \mathcal{S}(\Omega).$$

Equivalently, $\mathcal{S}(X^2)(N, \theta)$ *is the space of* $\varphi \in \mathcal{S}(X^2)$ *such that* $\varphi|_\Omega = 0$.

Proof. See Lemma 2.3 of [13]. ∎

Let $(X, \langle\cdot,\cdot\rangle)$ be a non-degenerate symmetric bilinear space over F satisfying (50), and let S be as in (11) with $\det(S) \neq 0$. Let $\Omega = \Omega_S$ be as in (61), and assume that Ω is non-empty. In Lemma 4.2.1 we noted that the subgroup B of $\mathrm{GL}(2, F) \times \mathrm{GO}(X)$ acts on Ω. By identifying $\mathrm{O}(X)$ with $1 \times \mathrm{O}(X) \subset \mathrm{GL}(2, F) \times \mathrm{GO}(X)$, we obtain an action of $\mathrm{O}(X)$ on Ω: this is given by $h \cdot (x_1, x_2) = (hx_1, hx_2)$, where $h \in \mathrm{O}(X)$ and $(x_1, x_2) \in \Omega$. This action is transitive. We obtain an action of $\mathrm{O}(X)$ on $\mathcal{S}(\Omega)$ by defining $(h \cdot \varphi)(x) = \varphi(h^{-1} \cdot x)$ for $h \in \mathrm{O}(X)$, $\varphi \in \mathcal{S}(\Omega)$ and $x \in \Omega$. This action is used in the next lemma.

4.4.5 Lemma. *Let $(X, \langle\cdot,\cdot\rangle)$ be a non-degenerate symmetric bilinear space over F satisfying (50), and let S be as in (11) with $\det(S) \neq 0$. Let $\Omega = \Omega_S$ be as in (61), and assume that Ω is non-empty. Let (σ_0, W_0) be an admissible representation of $\mathrm{O}(X)$, and let $M' : \mathcal{S}(\Omega) \to W_0$ be a non-zero $\mathrm{O}(X)$ map. Let $z \in \Omega$. There exists a compact, open subset C of Ω containing z such that if C_0 is a compact, open subset of Ω such that $z \in C_0 \subset C$, then $M'(f_{C_0}) \neq 0$. Here, f_{C_0} is the characteristic function of C_0.*

Proof. Let H be the subgroup of $h \in \mathrm{O}(X)$ such that $hz = z$. By 1.6 of [3], the map $H\backslash \mathrm{O}(X) \xrightarrow{\sim} \Omega$ defined by $Hh \mapsto h^{-1}z$ is a homeomorphism, so that the map $\mathcal{S}(\Omega) \xrightarrow{\sim} \text{c-Ind}_H^{\mathrm{O}(X)} 1_H$ that sends φ to the function f such that $f(h) = \varphi(h^{-1}z)$ for $h \in \mathrm{O}(X)$ is an $\mathrm{O}(X)$ isomorphism. Via this isomorphism, we may regard M' as defined on $\text{c-Ind}_H^{\mathrm{O}(X)} 1_H$, and it will suffice to prove that that there exists a compact, open neighborhood C of the identity in $\mathrm{O}(X)$ such that if C_0 is a compact, open neighborhood of the identity in $\mathrm{O}(X)$ with $C_0 \subset C$, then $M'(f_{HC_0}) \neq 0$, where f_{HC_0} is the characteristic function of HC_0. Since σ_0 is admissible, by 2.15 of [3] we have $(\sigma_1)^\vee \cong \sigma_0$ where $\sigma_1 = \sigma_0^\vee$. Let W_1 be the space of σ_1. We may regard M' as a non-zero element of $\mathrm{Hom}_{\mathrm{O}(X)}(\text{c-Ind}_H^{\mathrm{O}(X)} 1_H, \sigma_1^\vee)$. Now H and $\mathrm{O}(X)$ are unimodular since both are orthogonal groups (H is isomorphic to $\mathrm{O}(U^\perp)$, where $U = Fz_1 + Fz_2$). By 2.29 of [3], there exists an element λ of $\mathrm{Hom}_H(\sigma_1, 1_H)$ such that M' is given by

$$M'(f)(v) = \int\limits_{H\backslash \mathrm{O}(X)} f(h)\lambda(\sigma_1(h)v)\, dh$$

for $f \in \text{c-Ind}_H^{\mathrm{O}(X)} 1_H$ and $v \in W_1$. Since M' is non-zero, there exists $v \in W_1$ such that $\lambda(v) \neq 0$. Let C be a compact, open neighborhood of 1 in $\mathrm{O}(X)$ such that $\sigma_1(h)v = v$ for $h \in C$. Let C_0 be a compact, open neighborhood of 1 in $\mathrm{O}(X)$ such that $C_0 \subset C$. Then

$$\begin{aligned} M'(f_{HC_0})(v) &= \int\limits_{H\backslash \mathrm{O}(X)} f_{HC_0}(h)\lambda(\sigma_1(h)v)\, dh \\ &= \int\limits_{H\backslash HC_0} \lambda(\sigma_1(h)v)\, dh \\ &= \mathrm{vol}(H\backslash HC_0)\lambda(v), \end{aligned}$$

which is non-zero. ∎

In the following theorem we mention the set $\mathcal{E}(z)$ of embeddings of T into $\mathrm{GSO}(X)$; see (63), (67) and (69).

4.4.6 Theorem. *Let $(X, \langle\cdot,\cdot\rangle)$ be a non-degenerate symmetric bilinear space over F satisfying (50), and let S be as in (11) with $\det(S) \neq 0$. Let $A = A_S$,*

$T = T_S$, and $L = L_S$ be as in Sect. 2.3. If $\dim X = 4$, assume that A is a field. Let (π, V) be an irreducible, admissible representation of $\mathrm{GSp}(4,F)^+$, and let (σ, W) be an irreducible, admissible representation of $\mathrm{GO}(X)$. Assume that there is a non-zero R map $M : \mathcal{S}(X^2) \to \pi \otimes \sigma$. Let $\theta = \theta_S$ and let Λ be a character of T.

i) If $\mathrm{Hom}_N(\pi, \mathbb{C}_\theta) \neq 0$, then $\Omega = \Omega_S$ is non-empty and $D = TN$ is contained in $\mathrm{GSp}(4,F)^+$.

ii) Assume that $\mathrm{Hom}_N(\pi, \mathbb{C}_\theta) \neq 0$ so that $\Omega = \Omega_S$ is non-empty, and $D = TN \subset \mathrm{GSp}(4,F)^+$ by i). Assume further that $\mathrm{Hom}_D(\pi, \mathbb{C}_{\Lambda\otimes\theta}) \neq 0$. Let $z \in \Omega$, and $\tau \in \mathcal{E}(z)$. There exists a non-zero vector $w \in W$ such that

$$\sigma(\tau(t))w = \Lambda^{-1}(t)w$$

for $t \in T$.

Proof. i) The assumptions $\mathrm{Hom}_R(\mathcal{S}(X^2), V \otimes W) \neq 0$ and $\mathrm{Hom}_N(V, \mathbb{C}_\theta) \neq 0$ imply that $\mathrm{Hom}_N(\mathcal{S}(X^2), \mathbb{C}_\theta) \neq 0$. This means that $\mathcal{S}(X^2)_{N,\theta} \neq 0$; by Lemma 4.4.4, we obtain $\Omega \neq \varnothing$. Lemma 4.4.2 now also yields that $D \subset \mathrm{GSp}(4,F)^+$.
ii) Let β be a non-zero element of $\mathrm{Hom}_D(\pi, \mathbb{C}_{\Lambda\otimes\theta})$. We first claim that the composition M'

$$\mathcal{S}(X^2) \xrightarrow{M} V \otimes W \xrightarrow{\beta\otimes\mathrm{id}} \mathbb{C}_{\Lambda\otimes\theta} \otimes W$$

is non-zero. Let $\varphi \in \mathcal{S}(X^2)$ be such that $M(\varphi) \neq 0$, and write

$$M(\varphi) = \sum_{\ell=1}^{t} v_\ell \otimes w_\ell$$

where $v_1, \dots, v_t \in V$ and $w_1, \dots, w_t \in W$. We may assume that the vectors $w_1, \dots, w_t$ are linearly independent and that $v_1 \neq 0$. Since β is non-zero and V is an irreducible representation of $\mathrm{GSp}(4,F)^+$, it follows that there exists $g \in \mathrm{GSp}(4,F)^+$ such that $\beta(\pi(g)v_1) \neq 0$. Let $h \in \mathrm{GO}(X)$ be such that $\lambda(h) = \lambda(g)$. Then $(g,h) \in R$. Since M is an R-map, we have

$$M(\omega(g,h)\varphi) = \sum_{\ell=1}^{t} \pi(g)v_\ell \otimes \sigma(h)w_\ell.$$

Applying $\beta \otimes \mathrm{id}$ to this equation, we get

$$M'(\omega(g,h)\varphi) = \sum_{\ell=1}^{t} \beta(\pi(g)v_\ell) \otimes \sigma(h)w_\ell$$

in $\mathbb{C}_{\Lambda\otimes\theta} \otimes W$. Since the vectors $\sigma(h)w_1, \dots, \sigma(h)w_t$ are also linearly independent, and since $\beta(\pi(g)v_1)$ is non-zero, it follows that the vector $M'(\omega(g,h)\varphi)$ is non-zero; this proves $M' \neq 0$.

Next, the map M' induces a non-zero map $\mathcal{S}(X^2)_{N,\theta} \to \mathbb{C}_{\Lambda\otimes\theta} \otimes W$, which we also denote by M'. Lemma 4.4.4 implies that the restriction map yields an isomorphism $\mathcal{S}(X^2)_{N,\theta} \xrightarrow{\sim} \mathcal{S}(\Omega)$. Composing, we thus obtain a non-zero map $\mathcal{S}(\Omega) \to \mathbb{C}_{\Lambda\otimes\theta} \otimes W$, which we again denote by M'. Let $z \in \Omega$ and $\tau \in \mathcal{E}(z)$. By Lemma 4.2.1, the elements $([\begin{smallmatrix} & 1\\ 1 & \end{smallmatrix}]\, t\, [\begin{smallmatrix} & 1\\ 1 & \end{smallmatrix}], \tau(t))$ for $t \in T$ act on Ω. We can regard these elements as acting on $\mathcal{S}(\Omega)$ via the definition $\big(([\begin{smallmatrix} & 1\\ 1 & \end{smallmatrix}]\, t\, [\begin{smallmatrix} & 1\\ 1 & \end{smallmatrix}], \tau(t))\cdot\varphi\big)(x) = \varphi(([\begin{smallmatrix} & 1\\ 1 & \end{smallmatrix}]\, t\, [\begin{smallmatrix} & 1\\ 1 & \end{smallmatrix}], \tau(t))^{-1}\cdot x)$ for $\varphi \in \mathcal{S}(\Omega)$ and $x \in \Omega$. Moreover, by the definition of M' and (80), we have

$$M'(([\begin{smallmatrix} & 1\\ 1 & \end{smallmatrix}]\, t\, [\begin{smallmatrix} & 1\\ 1 & \end{smallmatrix}], \tau(t)) \cdot \varphi) = (\det(t), \mathrm{disc}(X))_F \Lambda(t)\sigma(\tau(t))M'(\varphi) \tag{84}$$

for $t \in T$ and $\varphi \in \mathcal{S}(\Omega)$. Let C be the compact, open subset from Lemma 4.4.5 with respect to M' and z; note that the restriction of σ to $\mathrm{O}(X)$ is admissible. By Lemma 4.2.5 there exists a compact, open subset C_0 of C containing z such that $([\begin{smallmatrix} & 1\\ 1 & \end{smallmatrix}]\, t\, [\begin{smallmatrix} & 1\\ 1 & \end{smallmatrix}], \tau(t)) \cdot C_0 = C_0$ for $t \in T$. Let $\varphi = f_{C_0}$. Then $([\begin{smallmatrix} & 1\\ 1 & \end{smallmatrix}]\, t\, [\begin{smallmatrix} & 1\\ 1 & \end{smallmatrix}], \tau(t)) \cdot \varphi = \varphi$ for $t \in T$, and by Lemma 4.4.5, we have $M'(\varphi) \neq 0$. From (84) we have $\sigma(\tau(t))M'(\varphi) = (\det(t), \mathrm{disc}(X))_F \Lambda(t)^{-1} M'(\varphi) = \chi_{L/F}(\mathrm{N}_{L/F}(t))\Lambda(t)^{-1}M'(\varphi) = \Lambda(t)^{-1}M'(\varphi)$ for $t \in T$. Since $M'(\varphi) \neq 0$, this proves ii). ∎

Let $(X, \langle\cdot,\cdot\rangle)$ be a non-degenerate symmetric bilinear space over F satisfying (50), and let S be as in (11) with $\det(S) \neq 0$. If Ω_S is non-empty and $z = (z_1, z_2) \in \Omega_S$, then we let $\mathrm{O}(X)_z$ be the subgroup of $h \in \mathrm{O}(X)$ such that $h(z_1) = z_1$ and $h(z_2) = z_2$.

4.4.7 Proposition. *Let $(X, \langle\cdot,\cdot\rangle)$ be a non-degenerate symmetric bilinear space over F satisfying (50), and assume that $\dim X = 4$. Let S be as in (11) with $\det(S) \neq 0$. Assume that Ω_S is non-empty, and let z be in Ω_S. Let $\varPi$ and σ be irreducible, admissible, supercuspidal representations of $\mathrm{GSp}(4, F)$ and $\mathrm{GO}(X)$, respectively. If $\mathrm{Hom}_R(\omega, \varPi \otimes \sigma) \neq 0$, then*

$$\dim \varPi_{N,\theta_S} = \dim \mathrm{Hom}_{\mathrm{O}(X)_z}(\sigma, \mathbb{C}_1). \tag{85}$$

Proof. Assume that $\mathrm{Hom}_R(\omega, \varPi \otimes \sigma) \neq 0$. By Proposition 3.3 of [27] the restriction of σ to $\mathrm{O}(X)$ is multiplicity-free. By Lemma 4.2 of [24] we have $\varPi|_{\mathrm{Sp}(4,F)} = \varPi_1 \oplus \cdots \oplus \varPi_t$, where $\varPi_1, \ldots, \varPi_t$ are mutually non-isomorphic, irreducible, admissible representations of $\mathrm{Sp}(4, F)$, $\sigma|_{\mathrm{O}(X)} = \sigma_1 \oplus \cdots \oplus \sigma_t$, where $\sigma_1, \ldots, \sigma_t$ are mutually non-isomorphic, irreducible, admissible representations of $\mathrm{O}(X)$, with $\mathrm{Hom}_{\mathrm{Sp}(4,F)\times\mathrm{O}(X)}(\omega, \varPi_i \otimes \sigma_i) \neq 0$ for $i \in \{1, \ldots, t\}$. Let $i \in \{1, \ldots, t\}$; to prove the proposition, it will suffice to prove that $(\varPi_i)_{N,\theta_S} \cong \mathrm{Hom}_{\mathrm{O}(X)_z}(\sigma_i, \mathbb{C}_1)$ as complex vector spaces. By Lemma 6.1 of [26], we have $\Theta(\sigma_i)_{N,\theta_S} \cong \mathrm{Hom}_{\mathrm{O}(X)_z}(\sigma_i^\vee, \mathbb{C}_1)$ as complex vector spaces. By 1) a) of the theorem on p. 69 of [14], the representation $\Theta(\sigma_i)$ of $\mathrm{Sp}(4, F)$ is irreducible. By Theorem 2.1 of [12] we have $\varPi_i \cong \Theta(\sigma_i)$. Therefore, $(\varPi_i)_{N,\theta_S} \cong \mathrm{Hom}_{\mathrm{O}(X)_z}(\sigma_i^\vee, \mathbb{C}_1)$. By the first theorem on p. 91 of [14], $\sigma_i^\vee \cong \sigma_i$. The proposition follows. ∎

4.5 Representations of $\mathrm{GO}(X)$

Let $m, \lambda \in F^\times$. By Lemma 4.1.1, the group $\mathrm{GSO}(X_{m,\lambda})$ is abelian. It follows that the irreducible, admissible representations of $\mathrm{GSO}(X_{m,\lambda})$ are characters. To describe the representations of $\mathrm{GO}(X_{m,\lambda})$, let $\mu : \mathrm{GSO}(X_{m,\lambda}) \to \mathbb{C}^\times$ be a character. We recall that the map γ from (53) is a representative for the non-trivial coset of $\mathrm{GSO}(X_{m,\lambda})$ in $\mathrm{GO}(X_{m,\lambda})$. Define $\mu^\gamma : \mathrm{GSO}(X_{m,\lambda}) \to \mathbb{C}^\times$ by $\mu^\gamma(x) = \mu(\gamma x \gamma^{-1})$. If $\mu^\gamma \neq \mu$, then the representation $\mathrm{ind}_{\mathrm{GSO}(X_{m,\lambda})}^{\mathrm{GO}(X_{m,\lambda})} \mu$ is irreducible, and we define

$$\mu^+ = \mathrm{ind}_{\mathrm{GSO}(X_{m,\lambda})}^{\mathrm{GO}(X_{m,\lambda})} \mu.$$

Assume that $\mu = \mu^\gamma$. Then the induced representation $\mathrm{ind}_{\mathrm{GSO}(X_{m,\lambda})}^{\mathrm{GO}(X_{m,\lambda})} \mu$ is reducible, and is the direct sum of the two extensions of μ to $\mathrm{GO}(X_{m,\lambda})$. We let μ^+ be the extension of μ to $\mathrm{GO}(X_{m,\lambda})$ such that $\mu^+(\gamma) = 1$ and let μ^- be the extension of μ to $\mathrm{GO}(X_{m,\lambda})$ such that $\mu^-(\gamma) = -1$. Every irreducible, admissible representation of $\mathrm{GO}(X_{m,\lambda})$ is of the form μ^+ or μ^- for some character μ of $\mathrm{GSO}(X_{m,\lambda})$. We will sometimes identify characters of $\mathrm{GSO}(X_{m,\lambda})$ with characters of $T_{\left[\begin{smallmatrix} 1 & \\ & -m \end{smallmatrix}\right]}$, via (54), and in turn identify characters of $T_{\left[\begin{smallmatrix} 1 & \\ & -m \end{smallmatrix}\right]}$ with characters of $L^\times$, via (14). Here L is associated to m, as in Sect. 2.1, so that $L = F(\sqrt{m})$ if $m \notin F^{\times 2}$, and $L = F \times F$ if $m \in F^{\times 2}$.
Next, let $(X, \langle \cdot, \cdot \rangle)$ be either $(X_{\mathrm{M}_2}, \langle \cdot, \cdot \rangle_{\mathrm{M}_2})$ or $(X_H, \langle \cdot, \cdot \rangle_H)$, as in (55) or (57). If $X = X_{\mathrm{M}_2}$, set $G = \mathrm{GL}(2, F)$, and if $X = X_H$, set $G = H^\times$. Let h_0 be the element of $\mathrm{GO}(X)$ that maps x to x^*; then h_0 represents the non-trivial coset of $\mathrm{GSO}(X)$ in $\mathrm{GO}(X)$. Let π_1 and π_2 be irreducible, admissible representations of G with the same central character. Via the exact sequences (58) and (59), the representations (π_1, V_1) and (π_2, V_2) define an irreducible, admissible representation $\pi_1 \otimes \pi_2$ of $\mathrm{GSO}(X)$ which has space $V_1 \otimes V_2$ and action given by the formula $(\pi_1 \otimes \pi_2)(\rho(g_1, g_2)) = \pi_1(g_1) \otimes \pi_2(g_2)$ for $g_1, g_2 \in G$. If π_1 and π_2 are not isomorphic, then $\pi_1 \otimes \pi_2$ induces irreducibly to $\mathrm{GO}(X)$; we denote this induced representation by $(\pi_1 \otimes \pi_2)^+$. Assume that π_1 and π_2 are isomorphic. In this case the representation $\pi_1 \otimes \pi_2$ does not induce irreducibly to $\mathrm{GO}(X)$, but instead has two extensions σ_1 and σ_2 to representations of $\mathrm{GO}(X)$. Moreover, the space of linear forms on $\pi_1 \otimes \pi_2$ that are invariant under the subgroup of $\mathrm{GSO}(X)$ of elements $\rho(g, g^{*-1})$ for $g \in G$ is one-dimensional. Let λ be a non-zero functional in this space. Then $\lambda \circ \sigma_i(h_0)$ is another such functional, so that $\lambda \circ \sigma_i(h_0) = \varepsilon_i \lambda$ with $\{\varepsilon_1, \varepsilon_2\} = \{1, -1\}$. The representation σ_i for which $\varepsilon_i = 1$ is denoted by $(\pi_1 \otimes \pi_2)^+$, and the representation σ_j for which $\varepsilon_j = -1$ is denoted by $(\pi_1 \otimes \pi_2)^-$. See [26] for details.

4.5.1 Proposition. *Let H be as in (56) and let X_H be as in (57). Let S be as in (72) with $-c \notin F^{\times 2}$; we may assume that $i^2 = -c$, as in (74). Let $z = (z_1, z_2)$ be as in (75), so that $z \in \Omega_S$. Set $L = F(\sqrt{-c})$. We have*

$$\dim \mathrm{Hom}_{\mathrm{O}(X_H)_z}(\sigma_0, \mathbb{C}_1) = 1$$

for the following families of irreducible, admissible representations σ_0 *of* $\mathrm{GO}(X_H)$*:*

i) $\sigma_0 = (\sigma 1_{H^\times} \otimes \sigma\chi_{L/F})^+$*;*

ii) $\sigma_0 = (\sigma 1_{H^\times} \otimes \sigma\pi^{\mathrm{JL}})^+$*.*

Here, σ *is a character of* $F^\times$*, and* π *is a supercuspidal, irreducible, admissible representation of* $\mathrm{GL}(2,F)$ *with trivial central character with* $\mathrm{Hom}_{L^\times}(\pi^{\mathrm{JL}},\mathbb{C}_1) \neq 0$.

Proof. We begin by describing $\mathrm{O}(X_H)_z$. Define $g_1 : X_H \to X_H$ by

$$g_1(1) = 1, \quad g_1(i) = i, \quad g_1(j) = j, \quad g_1(k) = -k.$$

Evidently, $g_1 \in \mathrm{O}(X_H)_z$, moreover, $\det(g_1) = -1$. It follows that $\mathrm{O}(X_H)_z = (\mathrm{SO}(X_H) \cap \mathrm{O}(X_H)_z) \sqcup (\mathrm{SO}(X_H) \cap \mathrm{O}(X_H)_z)g_1$. Using that $z_1 = 1$, $z_2 = i$, and the fact that every element of $\mathrm{SO}(X_H)$ is of the form $\rho(h_1,h_2)$ for some $h_1,h_2 \in H^\times$, a calculation shows that $\mathrm{SO}(X_H) \cap \mathrm{O}(X_H)_z$ is $\{\rho(h^{*-1},h) : h \in (F+Fi)^\times = L^\times\}$.
i) Since $\sigma_0|_{\mathrm{O}(X_H)} = (1_{H^\times} \otimes \chi_{L/F})^+$, we may assume that $\sigma = 1$. A model for σ_0 is $\mathbb{C} \oplus \mathbb{C}$, with action defined by

$$\sigma_0(\rho(h_1,h_2))(w_1 \oplus w_2) = \chi_{L/F}(\mathrm{N}(h_2))w_1 \oplus \chi_{L/F}(\mathrm{N}(h_1))w_2,$$
$$\sigma_0(*)(w_1 \oplus w_2) = w_2 \oplus w_1$$

for $w_1,w_2 \in \mathbb{C}$ and $h_1,h_2 \in H^\times$; here, $*$ is the canonical involution of H, regarded as an element of $\mathrm{O}(X_H)$ with determinant -1. Using that $g_1 = * \circ \rho(k^{*-1},k)$, we find that the restriction of σ_0 to $\mathrm{O}(X_H)_z$ is given by

$$\sigma_0(\rho(h^{*-1},h))(w_1 \oplus w_2) = w_1 \oplus w_2,$$
$$\sigma_0(g_1)(w_1 \oplus w_2) = \chi_{L/F}(\mathrm{N}(k))(w_2 \oplus w_1)$$

for $w_1,w_2 \in \mathbb{C}$ and $h \in (F+Fi)^\times = L^\times$. Therefore, $\sigma_0|_{\mathrm{O}(X_H)_z}$ is the direct sum of the trivial character $\mathrm{O}(X_H)_z$, and the non-trivial character of $\mathrm{O}(X_H)_z$ that is trivial on $\mathrm{SO}(X_H) \cap \mathrm{O}(X_H)_z$ and sends g_1 to -1. This implies that $\mathrm{Hom}_{\mathrm{O}(X_H)_z}(\sigma_0,\mathbb{C}_1)$ is one-dimensional.
ii) Again, we may assume that $\sigma = 1$. Let V be the space of π^{JL}. As a model for σ_0 we take $V \oplus V$ with action of $\mathrm{GO}(X_H)$ defined by

$$\sigma_0(\rho(h_1,h_2))(v_1 \oplus v_2) = \pi^{\mathrm{JL}}(h_2)v_1 \otimes \pi^{\mathrm{JL}}(h_1)v_2,$$
$$\sigma_0(*)(v_1 \oplus v_2) = v_2 \oplus v_1$$

for $h_1,h_2 \in H^\times$ and $v_1,v_2 \in V$. By hypothesis, $\mathrm{Hom}_{L^\times}(\pi^{\mathrm{JL}},\mathbb{C}_1) \neq 0$. This space is one-dimensional; see Sect. 2.7. We have $kLk^{-1} = L$; in fact, conjugation by k on L is the non-trivial element of $\mathrm{Gal}(L/F)$. Since $\mathrm{Hom}_{L^\times}(\pi^{\mathrm{JL}},\mathbb{C}_1)$

is one-dimensional, there exists $\varepsilon \in \{\pm 1\}$ such that $\lambda \circ \pi^{\mathrm{JL}}(k) = \varepsilon\lambda$ for $\lambda \in \mathrm{Hom}_{L^\times}(\pi^{\mathrm{JL}}, \mathbb{C}_1)$. Define a map

$$\mathrm{Hom}_{L^\times}(\pi^{\mathrm{JL}}, \mathbb{C}_1) \longrightarrow \mathrm{Hom}_{\mathrm{O}(X_H)_z}(\sigma_0, \mathbb{C}_1)$$

by sending λ to Λ, where Λ is defined by $\Lambda(v_1 \oplus v_2) = \lambda(v_1) + \varepsilon\lambda(v_2)$ for $v_1, v_2 \in V$. A computation using the fact that $g_1 = * \circ \rho(k^{*-1}, k)$ shows that this map is well defined. It is straightforward to verify that this map is injective and surjective, so that

$$\mathrm{Hom}_{L^\times}(\pi^{\mathrm{JL}}, \mathbb{C}_1) \cong \mathrm{Hom}_{\mathrm{O}(X_H)_z}(\sigma_0, \mathbb{C}_1).$$

Hence, $\mathrm{Hom}_{\mathrm{O}(X_H)_z}(\sigma_0, \mathbb{C}_1)$ is one-dimensional. ∎

4.6 GO(X) and GSp($4, F$)

In this section we will gather together some information about the theta correspondence between $\mathrm{GO}(X)$ and $\mathrm{GSp}(4)$ when X is as in (50). When $\dim(X) = 4$, we recall in Theorem 4.6.3 some results from [7] and [8]. When $\dim(X) = 2$, we calculate two theta lifts, producing representations of type Vd and IXb, in Proposition 4.6.2. This calculation uses P_3-theory. We include this material because, to the best of our knowledge, such a computation is absent from the literature.

We let R_Q be the group of elements of R of the form $(\begin{bmatrix} * & * & * & * \\ & * & * & * \\ & * & * & * \\ & & & * \end{bmatrix}, *)$. Let Z^J be the group defined in (40).

4.6.1 Lemma. *Let $(X, \langle\cdot,\cdot\rangle)$ be an even-dimensional symmetric bilinear space satisfying (50); assume additionally that X is anisotropic. There is an isomorphism of complex vector spaces*

$$T_1 : \mathcal{S}(X^2)_{Z^J} \xrightarrow{\ \sim\ } \mathcal{S}(X) \tag{86}$$

that is given by

$$T_1\big(\varphi + \mathcal{S}(X^2)(Z^J)\big)(x) = \varphi(x, 0)$$

for φ in $\mathcal{S}(X^2)$ and x in X. The subgroup R_Q of R acts on the quotient $\mathcal{S}(X^2)_{Z^J}$. Transferring this action to $\mathcal{S}(X)$ via T_1, the formulas for the resulting action are

$$(\begin{bmatrix} t & & & \\ & a & b & \\ & c & d & \\ & & & \lambda(h)t^{-1} \end{bmatrix}, h) \cdot \varphi = |\lambda(h)|^{-\dim(X)/4}\,(t, \mathrm{disc}(X))_F\,|t|^{\dim(X)/2}\,\omega_1([\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}], h)\varphi, \tag{87}$$

$$(\begin{bmatrix} 1 & x & y & z \\ & 1 & & y \\ & & 1 & -x \\ & & & 1 \end{bmatrix}, 1) \cdot \varphi = \varphi \tag{88}$$

for φ in $\mathcal{S}(X)$, x, y and z in F, t in $F^\times$, and $g = [\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}]$ in $\mathrm{GL}(2, F)$ and h in $\mathrm{GO}(X)$ with $\lambda(h) = \det(g)$.

Proof. We first claim that

$$\mathcal{S}(X^2)(Z^J) = \{\varphi \in \mathcal{S}(X^2) : \varphi(X \times 0) = 0\}. \tag{89}$$

Let φ be in $\mathcal{S}(X^2)(Z^J)$. By the lemma in 2.33 of [3] there exists a positive integer n so that

$$\int\limits_{\mathfrak{p}^{-n}} \omega(\begin{bmatrix} 1 & & b & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix}, 1)\varphi \, db = 0. \tag{90}$$

Evaluating at $(x, 0)$ and using (79) shows that $\varphi(X \times 0) = 0$. Conversely, assume that φ is contained in the right hand side of (89). For any integer k let

$$L_k = \{x \in X : \langle x, x\rangle \in \mathfrak{p}^k\}. \tag{91}$$

It is known that L_k is a lattice, i.e., it is a compact and open $\mathfrak{o}$ submodule of X; see the proof of Theorem 91:1 of [18]. Any lattice is free of rank $\dim X$ as an $\mathfrak{o}$ module. Since $\varphi(X \times 0) = 0$, there exists a positive integer n such that $\varphi(X \times L_n) = 0$. We claim that (90) holds. Let x_1 and x_2 be in X. Evaluating (90) at (x_1, x_2) gives

$$\Big(\int\limits_{\mathfrak{p}^{-n}} \psi(b\langle x_2, x_2\rangle) \, db\Big)\varphi(x_1, x_2).$$

This is zero if x_2 is in L_n because $\varphi(X \times L_n) = 0$. Assume that x_2 is not in L_n. By the definition of L_n, we have $\langle x_2, x_2\rangle \notin \mathfrak{p}^n$. This implies that

$$\int\limits_{\mathfrak{p}^{-n}} \psi(b\langle x_2, x_2\rangle) \, db = 0,$$

proving our claim. This completes the proof of (89).
Using (89), it is easy to verify that the map T_1 is an isomorphism of vector spaces. Equation (87) follows from Lemma 4.4.3, and equation (88) follows from (79) and (80). ∎

4.6.2 Proposition. *Let $m \in F^\times$, and let $(X_{m,1}, \langle\cdot,\cdot\rangle_{m,1})$ be as (52). Assume that $m \notin F^{\times 2}$, so that $X_{m,1}$ is anisotropic. Let $E = F(\sqrt{m})$, and identify characters of $\mathrm{GSO}(X_{m,1})$ and characters of $E^\times$ via (14) and (54). Let $\chi_{E/F}$ be the quadratic character associated to E. Let Π be an irreducible, admissible representation of $\mathrm{GSp}(4, F)$, and let σ be an irreducible, admissible representation of $\mathrm{GO}(X_{m,1})$.*

i) *Assume that $\sigma = \mu^+$ with $\mu = \mu \circ \gamma$, so that $\mu = \alpha \circ \mathrm{N}_{E/F}$ for a character α of $F^\times$. Then $\mathrm{Hom}_R(\omega, \Pi^\vee \otimes \sigma) \neq 0$ if and only if $\Pi = L(\nu\chi_{E/F}, \chi_{E/F} \rtimes \nu^{-1/2}\alpha)$ (type Vd).*

ii) Assume that $\sigma = \mu^+ = \mathrm{ind}_{\mathrm{GSO}(X_{m,1})}^{\mathrm{GO}(X_{m,1})}(\mu)$ *with* $\mu \neq \mu \circ \gamma$. *Then* $\mathrm{Hom}_R(\omega, \Pi^\vee \otimes \sigma) \neq 0$ *if and only if* $\Pi = L(\nu\chi_{E/F}, \nu^{-1/2}\pi(\mu))$ *(type IXb). Here,* $\pi(\mu)$ *is the supercuspidal, irreducible, admissible representation of* $\mathrm{GL}(2, F)$ *associated to* μ.

Proof. Let (σ, W) be as in i) or ii). In the case of i), set $\pi(\mu) = \alpha \times \alpha\chi_{E/F}$. Then $\mathrm{Hom}_{R_1}(\omega_1, \pi(\mu)^\vee \otimes \sigma) \neq 0$, and $\pi(\mu)$ is the unique irreducible, admissible representation of $\mathrm{GL}(2, F)$ with this property, by Theorem 4.6 of [11].
Let (Π', V) be an irreducible, admissible representation of $\mathrm{GSp}(4, F)$ such that $\mathrm{Hom}_R(\omega, \Pi' \otimes \sigma) \neq 0$. Let T be a non-zero element of this space. The non-vanishing of T implies that the central characters of Π' and σ satisfy

$$\omega_{\Pi'} = \omega_\sigma^{-1} = (\mu|_{F^\times})^{-1}. \tag{92}$$

We first claim that V is non-supercuspidal. By reasoning as in [9], there exist $\lambda_1, \ldots, \lambda_t$ in $F^\times$ and an irreducible $\mathrm{Sp}(4, F)$ subspace V_0 of V such that

$$V = V_1 \oplus \cdots \oplus V_t,$$

where

$$V_1 = \pi(\begin{bmatrix} 1 &&& \\ & 1 && \\ && \lambda_1 & \\ &&& \lambda_1 \end{bmatrix})V_0 \quad , \ldots , \quad V_t = \pi(\begin{bmatrix} 1 &&& \\ & 1 && \\ && \lambda_t & \\ &&& \lambda_t \end{bmatrix})V_0. \tag{93}$$

Similarly, there exist irreducible $\mathrm{O}(X)$ subspaces $W_1, \ldots, W_r$ of W such that

$$W = W_1 \oplus \cdots \oplus W_r.$$

There exists an i and a j such that $\mathrm{Hom}_{\mathrm{Sp}(4,F)\times\mathrm{O}(X)}(\omega, V_i \otimes W_j) \neq 0$. As in the proof of Lemma 4.2 of [24], there is an irreducible constituent U_1 of $\pi(\mu)^\vee$ such that $\mathrm{Hom}_{\mathrm{O}(X)}(\omega_1, U_1 \otimes W_j) \neq 0$. By Theorem 4.4 of [25], the representation V_i is non-supercuspidal, so that V is non-supercuspidal.
Since V is non-supercuspidal, we have $V_{Z^J} \neq 0$ by Tables A.5 and A.6 of [28] (see the comment after Theorem 3.2.1). We claim next that $\mathrm{Hom}_{R_Q}(\mathcal{S}(X^2)_{Z^J}, V_{Z^J} \otimes W) \neq 0$. It follows from (93) that $(V_i)_{Z^J} \neq 0$. Let $p_i : V \to V_i$ and $q_j : W \to W_j$ be the projections. These maps are $\mathrm{Sp}(4, F)$ and $\mathrm{O}(X)$ maps, respectively. The composition

$$\mathcal{S}(X^2) \xrightarrow{T} V \otimes W \xrightarrow{p_i \otimes q_j} V_i \otimes W_j \longrightarrow (V_i)_{Z^J} \otimes W_j$$

is non-zero and surjective; note that $V_i \otimes W_j$ is irreducible. The commutativity of the diagram

$$\begin{array}{ccccc} \mathcal{S}(X^2) & \xrightarrow{T} & V \otimes W & \xrightarrow{p_i \otimes q_j} & V_i \otimes W_j \\ & & \downarrow & & \downarrow \\ & & V_{Z^J} \otimes W & \xrightarrow{p_i \otimes q_j} & (V_i)_{Z^J} \otimes W_j \end{array}$$

implies our claim that $\mathrm{Hom}_{R_Q}(\mathcal{S}(X^2)_{Z^J}, V_{Z^J}\otimes W)\neq 0$.
Let $R_{\bar{Q}}$ be the subgroup of R_Q consisting of the elements of the form $(\begin{bmatrix} * & * & * & * \\ & * & * & * \\ & * & * & * \\ & & & 1 \end{bmatrix}, *)$. Let R_{P_3} be the subgroup of $P_3\times \mathrm{GO}(X)$ consisting of the elements of the form $(\begin{bmatrix} a & b & x \\ c & d & y \\ & & 1 \end{bmatrix}, h)$, $ad-bc=\lambda(h)$. There is a homomorphism from $R_{\bar{Q}}$ to R_{P_3} given by

$$(\begin{bmatrix} * & * & * & * \\ & a & b & x \\ & c & d & y \\ & & & 1 \end{bmatrix}, h)\mapsto (\begin{bmatrix} a & b & x \\ c & d & y \\ & & 1 \end{bmatrix}, h)$$

for $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ in $\mathrm{GL}(2,F)$, x and y in F, and h in $\mathrm{GO}(X)$ with $ad-bc=\lambda(h)$. We consider Z^J a subgroup of $R_{\bar{Q}}$ via $z\mapsto (z,1)$. The above homomorphism then induces an isomorphism $R_{\bar{Q}}/Z^J\cong R_{P_3}$.
We restrict the R_Q modules $\mathcal{S}(X^2)_{Z^J}$ and $V_{Z^J}\otimes W$ to $R_{\bar{Q}}$. The subgroup Z^J of $R_{\bar{Q}}$ acts trivially, so that these spaces may be viewed as R_{P_3} modules.
Let χ be a character of $F^\times$. We assert that

$$\mathrm{Hom}_{R_{P_3}}(\mathcal{S}(X^2)_{Z^J}, \tau^{P_3}_{\mathrm{GL}(0)}(1)\otimes\sigma)=0, \tag{94}$$

$$\mathrm{Hom}_{R_{P_3}}(\mathcal{S}(X^2)_{Z^J}, \tau^{P_3}_{\mathrm{GL}(1)}(\chi)\otimes\sigma)=0. \tag{95}$$

Let τ be $\tau^{P_3}_{\mathrm{GL}(0)}(1)$ or $\tau^{P_3}_{\mathrm{GL}(1)}(\chi)$. Assume that (94) or (95) is non-zero; we will obtain a contradiction. Let S be a non-zero element of (94) or (95). Since S is non-zero, there exists φ in $\mathcal{S}(X^2)_{Z^J}$ such that $S(\varphi)$ is non-zero. Write $S(\varphi)=\sum_{i=1}^t f_i\otimes w_i$ for some $f_1,\dots,f_t$ in the standard space of τ and $w_1,\dots,w_t$ in W. The elements $f_1,\dots,f_t$ are functions from P_3 to $\mathbb{C}$ such that

$$f_i(\begin{bmatrix} 1 & & x \\ & 1 & y \\ & & 1 \end{bmatrix} p)=\psi(y)f_i(p)$$

for x and y in F, p in P_3, and $i=1,\dots,t$. We may assume that the vectors $w_1,\dots,w_t$ are linearly independent, and that there exists p in P_3 such that $f_1(p)$ is non-zero. Using the transformation properties of S and f_1, we may assume that $p=\begin{bmatrix} a & & \\ & 1 & \\ & & 1 \end{bmatrix}$. Let $\lambda:\sigma\to\mathbb{C}$ be a linear functional such that $\lambda(w_1)=1$ and $\lambda(w_2)=\cdots=\lambda(w_t)=0$, and let $e:\tau\to\mathbb{C}$ be the linear functional that sends f to $f(p)$. The composition $(e\otimes\lambda)\circ S$ is non-zero on φ. On the other hand, using (88), for y in F we have

$$\begin{aligned}
\big((e\otimes\lambda)\circ S\big)((\begin{bmatrix} 1 & & \\ & 1 & y \\ & & 1 \end{bmatrix},1)\varphi) &= (e\otimes\lambda)\big((\begin{bmatrix} 1 & & \\ & 1 & y \\ & & 1 \end{bmatrix},1)\cdot S(\varphi)\big),\\
\big((e\otimes\lambda)\circ S\big)(\varphi) &= (e\otimes\lambda)\big((\begin{bmatrix} 1 & & \\ & 1 & y \\ & & 1 \end{bmatrix},1)\cdot \sum_{i=1}^t f_i\otimes w_i\big)\\
&= \sum_{i=1}^t f_i(p\begin{bmatrix} 1 & & \\ & 1 & y \\ & & 1 \end{bmatrix})\lambda(w_i)\\
&= \psi(y)f_1(p),
\end{aligned}$$

$$\big((e \otimes \lambda) \circ S\big)(\varphi) = \psi(y)\big((e \otimes \lambda) \circ S\big)(\varphi).$$

This is a contradiction since $\big((e \otimes \lambda) \circ S\big)(\varphi)$ is non-zero, and there exist y in F such that $\psi(y) \neq 1$. This concludes the proof of (94) and (95).

It follows from (94), (95) and the non-vanishing of $\mathrm{Hom}_{R_{P_3}}(\mathcal{S}(X^2)_{Z^J}, V_{Z^J} \otimes W)$ that there exists an irreducible, admissible representation ρ of $\mathrm{GL}(2,F)$ that occurs in the P_3 filtration of V_{Z^J} (see Theorem 3.2.1) such that $\mathrm{Hom}_{R_{P_3}}(\mathcal{S}(X^2)_{Z^J}, \tau^{P_3}_{\mathrm{GL}(2)}(\rho) \otimes W) \neq 0$. It follows from (87) that $\mathrm{Hom}_{R_1}(\omega_1, \nu^{-1/2}\chi_{E/F}\rho \otimes \sigma) \neq 0$. By the uniqueness stated in the first paragraph of this proof, it follows that

$$\rho = \nu^{1/2}\chi_{E/F}\,\pi(\mu)^\vee. \tag{96}$$

As a consequence, $\omega_\rho = \nu(\mu|_{F^\times})^{-1}\chi_{E/F}$. Together with (92), it follows that

$$\omega_{\Pi'} = \chi_{E/F}\nu^{-1}\omega_\rho. \tag{97}$$

Going through Table A.5 of [28], we see that only the $\Pi' = \Pi^\vee$ with Π as asserted in i) and ii) satisfy both (96) and (97). (Observe the remark made after Theorem 3.2.1.)

Conversely, assume that Π is as in i) or ii). Since $\mathrm{Hom}_{R_1}(\omega_1, \pi(\mu)^\vee \otimes \sigma) \neq 0$, we have $\mathrm{Hom}_{\mathrm{O}(X)}(\mathcal{S}(X^2), \sigma) \neq 0$ by, for example, Remarque b) on p. 67 of [14]. Arguing as in Theorem 4.4 of [24], there exists some irreducible, admissible representation Π' of $\mathrm{GSp}(4,F)$ such that $\mathrm{Hom}_R(\omega, \Pi' \otimes \sigma) \neq 0$. By what we proved above, $\Pi' = \Pi^\vee$. This concludes the proof. ∎

4.6.3 Theorem. ([7],[8]) *Let $(X, \langle\cdot,\cdot\rangle)$ be either $(X_{\mathrm{M}_2}, \langle\cdot,\cdot\rangle_{\mathrm{M}_2})$ or $(X_H, \langle\cdot,\cdot\rangle_H)$, as in (55) or (57). If $X = X_{\mathrm{M}_2}$, set $G = \mathrm{GL}(2,F)$, and if $X = X_H$, set $G = H^\times$. Let Π be an irreducible, admissible representation of $\mathrm{GSp}(4,F)$, and let π_1 and π_2 be irreducible, admissible representations of G with the same central character. We have*

$$\mathrm{Hom}_R(\omega, \Pi^\vee \otimes (\pi_1 \otimes \pi_2)^+) \neq 0$$

for Π, π_1 and π_2 as in the following table:

type of Π		Π	π_1	π_2
I		$\chi_1 \times \chi_2 \rtimes \sigma$	$\sigma\chi_1\chi_2 \times \sigma$	$\sigma\chi_1 \times \sigma\chi_2$
II	a	$\chi \mathrm{St}_{\mathrm{GL}(2)} \rtimes \sigma$	$\sigma\chi^2 \times \sigma$	$\sigma\chi\mathrm{St}_{\mathrm{GL}(2)}$
	b	$\chi 1_{\mathrm{GL}(2)} \rtimes \sigma$	$\sigma\chi^2 \times \sigma$	$\sigma\chi 1_{\mathrm{GL}(2)}$
III	b	$\chi \rtimes \sigma 1_{\mathrm{GSp}(2)}$	$\sigma\chi\nu^{1/2} \times \sigma\nu^{-1/2}$	$\sigma\chi\nu^{-1/2} \times \sigma\nu^{1/2}$
IV	c	$L(\nu^{3/2}\mathrm{St}_{\mathrm{GL}(2)}, \nu^{-3/2}\sigma)$	$\sigma\nu^{3/2} \times \sigma\nu^{-3/2}$	$\sigma\mathrm{St}_{\mathrm{GL}(2)}$
	d	$\sigma 1_{\mathrm{GSp}(4)}$	$\sigma\nu^{3/2} \times \sigma\nu^{-3/2}$	$\sigma 1_{\mathrm{GL}(2)}$
V	a	$\delta([\xi, \nu\xi], \nu^{-1/2}\sigma)$	$\sigma\mathrm{St}_{\mathrm{GL}(2)}$	$\sigma\xi\mathrm{St}_{\mathrm{GL}(2)}$
	a*	$\delta^*([\xi, \nu\xi], \nu^{-1/2}\sigma)$	$\sigma 1_{H^\times}$	$\sigma\xi 1_{H^\times}$
	b	$L(\nu^{1/2}\xi\mathrm{St}_{\mathrm{GL}(2)}, \nu^{-1/2}\sigma)$	$\sigma 1_{\mathrm{GL}(2)}$	$\sigma\xi\mathrm{St}_{\mathrm{GL}(2)}$
	d	$L(\nu\xi \rtimes \nu^{-1/2}\sigma)$	$\sigma 1_{\mathrm{GL}(2)}$	$\sigma\xi 1_{\mathrm{GL}(2)}$
VI	a	$\tau(S, \nu^{-1/2}\sigma)$	$\sigma\mathrm{St}_{\mathrm{GL}(2)}$	$\sigma\mathrm{St}_{\mathrm{GL}(2)}$
	b	$\tau(T, \nu^{-1/2}\sigma)$	$\sigma 1_{H^\times}$	$\sigma 1_{H^\times}$
	c	$L(\nu^{1/2}\mathrm{St}_{\mathrm{GL}(2)}, \nu^{-1/2}\sigma)$	$\sigma 1_{\mathrm{GL}(2)}$	$\sigma\mathrm{St}_{\mathrm{GL}(2)}$
	d	$L(\nu, 1_{F^\times} \rtimes \nu^{-1/2}\sigma)$	$\sigma 1_{\mathrm{GL}(2)}$	$\sigma 1_{\mathrm{GL}(2)}$
VIII	a	$\tau(S, \pi)$	π	π
	b	$\tau(T, \pi)$	π^{JL}	π^{JL}
X		$\pi \rtimes \sigma$	$\sigma\omega_\pi \times \sigma$	π
XI	a	$\delta(\nu^{1/2}\pi, \nu^{-1/2}\sigma)$	$\sigma\mathrm{St}_{\mathrm{GL}(2)}$	$\sigma\pi$
	a*	$\delta^*(\nu^{1/2}\pi, \nu^{-1/2}\sigma)$	$\sigma 1_{H^\times}$	$\sigma\pi^{\mathrm{JL}}$
	b	$L(\nu^{1/2}\pi, \nu^{-1/2}\sigma)$	$\sigma 1_{\mathrm{GL}(2)}$	$\sigma\pi$

The notation π^{JL} in the table denotes the Jacquet-Langlands lifting of the supercuspidal representation π of $\mathrm{GL}(2, F)$ to a representation of $H^\times$. See Sect. 4.5 for the definitions of the + representation.

4.7 Applications

We now apply Theorem 4.4.6 along with knowledge of the theta correspondences of the previous section to obtain results about Bessel functionals.

4.7.1 Corollary. *Let $(X, \langle\cdot,\cdot\rangle)$ be either $(X_{\mathrm{M}_2}, \langle\cdot,\cdot\rangle_{\mathrm{M}_2})$ or $(X_H, \langle\cdot,\cdot\rangle_H)$, as in (55) or (57). If $X = X_{\mathrm{M}_2}$, set $G = \mathrm{GL}(2, F)$, and if $X = X_H$, set $G = H^\times$.*

Let Π be an irreducible, admissible representation of $\mathrm{GSp}(4,F)$*, and let* π_1 *and* π_2 *be irreducible, admissible representations of* G *with the same central character. Assume that*

$$\mathrm{Hom}_R(\omega, \Pi^\vee \otimes (\pi_1 \otimes \pi_2)^+) \neq 0$$

and that Π *has a non-split* (Λ,θ)*-Bessel functional with* $\theta = \theta_S$*. Then*

$$\mathrm{Hom}_T(\pi_1, \mathbb{C}_\Lambda) \neq 0 \quad \textit{and} \quad \mathrm{Hom}_T(\pi_2, \mathbb{C}_\Lambda) \neq 0$$

where $T = T_S$*.*

Proof. The assumption that the Bessel functional is non-split means that $A = A_S$ is a field. By Sect. 2.2 and Sect. 2.5 we may assume that S has the diagonal form (72). By (19), the contragredient $\Pi^\vee$ has a $((\Lambda\circ\gamma)^{-1},\theta)$-Bessel functional. The assertion follows now from Theorem 4.4.6, the explicit embeddings in (73) and (76), and the relation (31). ∎

4.7.2 Corollary. *Let* (Π, V) *be an irreducible, admissible representation of* $\mathrm{GSp}(4,F)$*. If* Π *is one of the representations in the following table, then* Π *admits a non-zero* (Λ,θ)*-Bessel functional* β *if and only if the quadratic extension* L *associated to* β*, and* Λ*, regarded as a character of* $L^\times$*, are as specified in the table.*

type of Π	Π	L	Λ
Va*	$\delta^*([\chi_{E/F},\nu\chi_{E/F}],\nu^{-1/2}\alpha)$	E	$\alpha\circ \mathrm{N}_{E/F}$
Vd	$L(\nu\chi_{E/F},\chi_{E/F}\rtimes\nu^{-1/2}\alpha)$	E	$\alpha\circ \mathrm{N}_{E/F}$
IXb	$L(\nu\chi_{E/F},\nu^{-1/2}\pi(\mu))$	E	μ *and the Galois conjugate of* μ

Proof. First we consider the Va* case. Let $\Pi = \delta^*([\chi_{E/F},\nu\chi_{E/F}],\nu^{-1/2}\alpha)$. By Theorem 4.6.3,

$$\mathrm{Hom}_R(\omega, \Pi^\vee \otimes (\alpha 1_{H^\times} \otimes \alpha\chi_{E/F}1_{H^\times})^+) \neq 0.$$

First, assume that Π admits a non-zero (Λ,θ)-Bessel functional, and let L be the quadratic extension associated to this Bessel functional; we will prove that $E = L$ and that $\Lambda = \alpha\circ\mathrm{N}_{E/F}$. By v) of Proposition 3.5.1, this Bessel functional is non-split. It follows from Corollary 4.7.1 that

$$\alpha(\mathrm{N}_{L/F}(t)) = \Lambda(t) \qquad \text{and} \qquad (\chi_{E/F}\alpha)(\mathrm{N}_{L/F}(t)) = \Lambda(t)$$

for t in $T = L^\times$. It follows that $E = L$, and that $\Lambda = \alpha\circ\mathrm{N}_{E/F}$.
Finally, we prove that Va* admits a Bessel functional as specified in the statement of the corollary. By Theorem 6.1.4 below, Va* admits some non-zero

Bessel functional. The previous paragraph proves that this Bessel functional must be as described in the statement of the corollary.

The arguments for the cases Vd and IXb are similar; we will only consider the case of type IXb. Let $\Pi = L(\nu\chi_{E/F}, \nu^{-1/2}\pi(\mu))$, where E is a quadratic extension of F, $\chi_{E/F}$ is the quadratic character associated to E/F, μ is a character of $E^\times$ that is not Galois invariant, and $\pi(\mu)$ is the supercuspidal, irreducible, admissible representation of $\mathrm{GL}(2,F)$ associated to μ.

First, assume that Π admits a non-zero (Λ,θ)-Bessel functional, and let L be the quadratic extension associated to this Bessel functional; we will prove that $E=L$, and that Λ is μ or the Galois conjugate of μ. Let S define θ, as in (15). By (19), $\Pi^\vee$ admits a non-zero $((\Lambda\circ\gamma)^{-1},\theta)$ Bessel functional β. Write $E=F(\sqrt{m})$ for some $m\in F^\times$. By Proposition 4.6.2 we have $\mathrm{Hom}_R(\omega,\Pi^\vee\otimes\mu^+)\neq 0$ with μ^+ as in this proposition. The involved symmetric bilinear space is $(X_{m,1},\langle\cdot,\cdot\rangle_{m,1})$. Let $\mathrm{GSp}(4,F)^+$ be defined with respect to $(X_{m,1},\langle\cdot,\cdot\rangle_{m,1})$ as in (77). By Lemma 4.4.1 the index of $\mathrm{GSp}(4,F)^+$ in $\mathrm{GSp}(4,F)$ is two. By Lemma 2.1 of [9], the restriction of $\Pi^\vee$ to $\mathrm{GSp}(4,F)^+$ is irreducible or the direct sum of two non-isomorphic irreducible, admissible representations of $\mathrm{GSp}(4,F)^+$; the non-vanishing of $\mathrm{Hom}_R(\omega,\Pi^\vee\otimes\mu^+)$ and Lemma 4.1 of [27] (with $m=2$ and $n=2$) imply that $V^\vee=V_1\oplus V_2$ with V_1 and V_2 irreducible $\mathrm{GSp}(4,F)^+$ subspaces of V. Moreover, for each $i\in\{1,2\}$, there exists $\lambda_i\in F^\times$ such that $\Pi(\left[\begin{smallmatrix}1&\\&\lambda_i\end{smallmatrix}\right])V_1=V_i$. Since $\mathrm{Hom}_R(\omega,V^\vee\otimes\mu^+)$ is non-zero, we may assume, after possibly renumbering, that $\mathrm{Hom}_R(\omega,V_1\otimes\mu^+)$ is non-zero. There exists $i\in\{1,2\}$ such that the restriction of β to V_i is non-zero. Let $\beta'=\begin{bmatrix}1&\\&\lambda_i^{-1}\end{bmatrix}\cdot\beta$. From Sect. 2.5 it follows that β' is a $((\Lambda\circ\gamma)^{-1},\theta')$ Bessel functional on $\Pi^\vee$ with θ' defined by $S'=\lambda_i^{-1}S$; also, the restriction of β' to V_1 is non-zero. We will now apply Theorem 4.4.6, with S' and V_1 playing the roles of S and π, respectively. By i) of this theorem we have that $\Omega_{S'}$ is non-empty; since S and S' have the same discriminant, Lemma 4.3.1 implies that $L=E$. Let $z\in\Omega_{S'}$ and $\tau\in\mathcal{E}(z)$. By ii) of Theorem 4.4.6, there exists a non-zero vector w in the space of μ^+ such that $\mu^+(\tau(t))w=(\Lambda\circ\gamma)(t)w$ for $t\in T_{S'}$. By Lemma 4.3.1 again, this implies that $\mu^+(\rho(x))w=\Lambda(x)w$ for $x\in L^\times$, or $\mu^+(\rho(\gamma(x)))w=\Lambda(x)w$ for $x\in L^\times$. Since $w\neq 0$, the definition of μ^+ now implies that $\Lambda=\mu$ or $\mu\circ\gamma$, as desired.

Finally, we prove that Π admits Bessel functionals as specified in the statement of the corollary. By Theorem 6.1.4 below, Π admits some non-zero Bessel functional. The previous paragraph proves that this Bessel functional must be as described in the statement of the corollary, and (29) implies that the Π admits both of the asserted Bessel functionals. ∎

The following result will imply uniqueness of Bessel functionals for representations of type Va* and XIa*.

4.7.3 Corollary. *Let σ be a character of $F^\times$. Let $c\in F^\times$ with $-c\notin F^{\times 2}$. Let S be as in (72) and set $L=F(\sqrt{-c})$.*

i) If $\xi = \chi_{L/F}$*, then* $\dim \delta^*([\xi, \nu\xi], \nu^{-1/2}\sigma)_{N,\theta_S} = 1$.

ii) If π *is an irreducible, admissible, supercuspidal representation of* $\mathrm{GL}(2,F)$ *with trivial central character such that* $\mathrm{Hom}_{L^\times}(\pi^{\mathrm{JL}}, \mathbb{C}_1) \neq 0$*, then* $\dim \delta^*(\nu^{1/2}\pi, \nu^{-1/2}\sigma)_{N,\theta_S} = 1$.

Proof. This follows from Proposition 4.4.7, Theorem 4.6.3, and Proposition 4.5.1; note that $\Pi^\vee|_N \cong \Pi|_N$ for irreducible, admissible representations Π of $\mathrm{GSp}(4,F)$ because $\Pi^\vee \cong \omega_\Pi^{-1}\Pi$. ∎

5 Twisted Jacquet modules of induced representations

Let (π, V) be an irreducible, admissible representation of $\mathrm{GSp}(4,F)$. In view of the isomorphism (20), understanding the possible Bessel functionals of π is equivalent to understanding the twisted Jacquet modules $V_{N,\theta}$ as T-modules. In this section, we will calculate the twisted Jacquet modules for representations induced from the Siegel and Klingen parabolic subgroup. This information will be used to determine the possible Bessel functionals for many of the non-supercuspidal representations of $\mathrm{GSp}(4,F)$; see Sect. 6.2.
The results of this section are similar to Proposition 2.1 and 2.3 of [23]. However, we prefer to redo the arguments, as those in [23] contain some inaccuracies.

5.1 Two useful lemmas

For a positive integer n let $\mathcal{S}(F^n)$ be the Schwartz space of F^n, meaning the space of locally constant, compactly supported functions $F^n \to \mathbb{C}$. As before, ψ is our fixed non-trivial character of F.
Let V be a complex vector space. Let $\mathcal{S}(F,V)$ be the space of compactly supported, locally constant functions from F to V. There is a canonical isomorphism $\mathcal{S}(F,V) \cong \mathcal{S}(F) \otimes V$. The functional on $\mathcal{S}(F)$ given by $f \mapsto \int_F f(x)\,dx$ gives rise to a linear map $\mathcal{S}(F) \otimes V \to V$, and hence to a linear map $\mathcal{S}(F,V) \to V$. We write this map as an integral

$$f \longmapsto \int\limits_F f(x)\,dx.$$

The following lemma will be frequently used when we calculate Jacquet modules in the subsequent sections.

5.1.1 Lemma. *Let* ρ *denote the action of* F *on* $\mathcal{S}(F,V)$ *by translation, i.e.,* $(\rho(x)f)(y) = f(x+y)$*. Let* ρ' *be the action of* F *on* $\mathcal{S}(F,V)$ *given by* $(\rho'(x)f)(y) = \psi(xy)f(y)$.

i) The map $f \mapsto \int_F f(x)\,dx$ *induces an isomorphism*

$$\mathcal{S}(F,V)/\langle f - \rho(x)f \,:\, x \in F\rangle \cong V.$$

ii) The map $f \mapsto \int_F \psi(-x)f(x)\,dx$ *induces an isomorphism*

$$\mathcal{S}(F,V)/\langle \psi(x)f - \rho(x)f \;:\; x \in F\rangle \cong V.$$

iii) The map $f \mapsto f(0)$ *induces an isomorphism*

$$\mathcal{S}(F,V)/\langle f - \rho'(x)f \;:\; x \in F\rangle \cong V.$$

Proof. By the Proposition in 1.18 of [3], every translation-invariant functional on $\mathcal{S}(F)$ is a multiple of the Haar measure $f \mapsto \int_F f(x)\,dx$. This proves i) in the case where $V = \mathbb{C}$. The general case follows from this case by tensoring the exact sequence

$$0 \longrightarrow \langle f - \rho(x)f \;:\; x \in F,\ f \in \mathcal{S}(F)\rangle \longrightarrow \mathcal{S}(F) \longrightarrow \mathbb{C} \longrightarrow 0$$

by V. Under the isomorphism $\mathcal{S}(F)\otimes V \cong \mathcal{S}(F,V)$, the space $\langle f - \rho(x)f \;:\; x \in F,\ f \in \mathcal{S}(F)\rangle \otimes V$ maps onto $\langle f - \rho(x)f \;:\; x \in F,\ f \in \mathcal{S}(F,V)\rangle$.
To prove ii), observe that there is an isomorphism

$$\begin{aligned}\mathcal{S}(F,V)/\langle f - \rho(x)f \;:\; x \in F,\ f \in \mathcal{S}(F,V)\rangle&\\ \longrightarrow \mathcal{S}(F,V)/\langle \psi(x)f - \rho(x)f \;:\; x \in F,\ f \in \mathcal{S}(F,V)\rangle&\end{aligned}$$

induced by the map $f \mapsto f'$, where $f'(x) = \psi(x)f(x)$. Hence ii) follows from i). Finally, iii) also follows from i), since the Fourier transform $f \mapsto \hat{f}$, where

$$\hat{f}(y) = \int\limits_F \psi(-uy)f(u)\,du,$$

intertwines the actions ρ and ρ' of F on $\mathcal{S}(F,V)$. ∎

5.1.2 Lemma. *Let G be an l-group, as in [3], and let H_1 and H_2 be closed subgroups of G. Assume that $G = H_1H_2$, and that for every compact subset K of G, there exists a compact subset K_2 of H_2 such that $K \subset H_1K_2$. Let (ρ, V) be a smooth representation of H_1. The map $r : \text{c-Ind}_{H_1}^G \rho \to \text{c-Ind}_{H_1\cap H_2}^{H_2}(\rho|_{H_1\cap H_2})$ defined by restriction of functions is a well-defined isomorphism of representations of H_2.*

Proof. This follows from straightforward verifications. ∎

5.2 Siegel induced representations

Let π be an admissible representation of $\mathrm{GL}(2,F)$, let σ be a character of $F^\times$, and let $\pi \rtimes \sigma$ be as defined in Sect. 1.2; see (9). In this section we will calculate the twisted Jacquet modules $(\pi \rtimes \sigma)_{N,\theta}$ for any non-degenerate character θ of N as a module of T. Lemma 5.2.2 below corrects an inaccuracy in Proposition 2.1 of [23]. Namely, Proposition 2.1 of [23] does not include ii) of our lemma.

5.2.1 Lemma. *Let σ be a character of $F^\times$, and π an admissible representation of $\mathrm{GL}(2,F)$. Let I be the standard space of the Siegel induced representation $\pi \rtimes \sigma$. There is a filtration of P-modules*

$$I^3 = 0 \subset I^2 \subset I^1 \subset I^0 = I.$$

with the quotients given as follows.

i) *$I^0/I^1 = \sigma_0$, where*

$$\sigma_0(\left[\begin{smallmatrix} A & * \\ & cA' \end{smallmatrix}\right]) = \sigma(c)\,|c^{-1}\det(A)|^{3/2}\,\pi(A)$$

for A in $\mathrm{GL}(2,F)$ and c in $F^\times$.

ii) *$I^1/I^2 = \text{c-Ind}^P_{\left[\begin{smallmatrix} * & * & * & * \\ & * & & * \\ & & * & * \\ & & & * \end{smallmatrix}\right]}\sigma_1$, where*

$$\sigma_1(\left[\begin{smallmatrix} t & * & y & * \\ & a & & * \\ & & d & * \\ & & & adt^{-1} \end{smallmatrix}\right]) = \sigma(ad)\,|a^{-1}t|^{3/2}\,\pi(\left[\begin{smallmatrix} t & y \\ & d \end{smallmatrix}\right])$$

for y in F and a,d,t in $F^\times$.

iii) *$I^2/I^3 = \text{c-Ind}^P_{\left[\begin{smallmatrix} * & * & & \\ * & * & & \\ & & * & * \\ & & * & * \end{smallmatrix}\right]}\sigma_2$, where*

$$\sigma_2(\left[\begin{smallmatrix} A & \\ & cA' \end{smallmatrix}\right]) = \sigma(c)\,|c\det(A)^{-1}|^{3/2}\,\pi(cA')$$

for A in $\mathrm{GL}(2,F)$ and c in $F^\times$.

Proof. This follows by going through the procedure of Sections 6.2 and 6.3 of [5]. ∎

5.2.2 Lemma. *Let σ be a character of $F^\times$, and let (π, V) be an admissible representation of $\mathrm{GL}(2,F)$. We assume that π admits a central character ω_π. Let I be the standard space of the Siegel induced representation $\pi \rtimes \sigma$. Let θ be the character of N defined in (15). Assume that θ is non-degenerate. Let L be the quadratic extension associated to S as in Sect. 2.3.*

i) *Assume that L is a field. Then $I_{N,\theta} \cong V$ with the action of T given by $\sigma\omega_\pi\pi'$. Here, π' is the representation of $\mathrm{GL}(2,F)$ on V given by $\pi'(g) = \pi(g')$. In particular, if π is irreducible, then the action of T is given by $\sigma\pi$.*

ii) *Assume that L is not a field; we may arrange that $S = \left[\begin{smallmatrix} & 1/2 \\ 1/2 & \end{smallmatrix}\right]$. Then there is a filtration*

$$0 \subset J_2 \subset J_1 = I_{N,\theta},$$

with vector space isomorphisms:

- $J_1/J_2 \cong V_{\left[\begin{smallmatrix}1&*\\&1\end{smallmatrix}\right],\psi} \oplus V_{\left[\begin{smallmatrix}1&*\\&1\end{smallmatrix}\right],\psi}$,
- $J_2 \cong V$.

The action of $T = T_S$ is given as follows,

$$\mathrm{diag}(a,b,a,b)(v_1 \oplus v_2) = \left|\frac{a}{b}\right|^{1/2}\sigma(ab)\omega_\pi(a)v_1 \oplus \left|\frac{a}{b}\right|^{-1/2}\sigma(ab)\omega_\pi(b)v_2,$$
$$\mathrm{diag}(a,b,a,b)v = \sigma(ab)\pi(\left[\begin{smallmatrix}a&\\&b\end{smallmatrix}\right])v,$$

for $a,b \in F^\times$, $v_1 \oplus v_2 \in J_1/J_2 \cong V_{\left[\begin{smallmatrix}1&*\\&1\end{smallmatrix}\right],\psi} \oplus V_{\left[\begin{smallmatrix}1&*\\&1\end{smallmatrix}\right],\psi}$, and $v \in J_2$. In particular, if π is one-dimensional, then $I_{N,\theta} \cong V$, with the action of T given by $\mathrm{diag}(a,b,a,b)v = \sigma(ab)\pi(\left[\begin{smallmatrix}a&\\&b\end{smallmatrix}\right])v$.

Proof. We may assume that $b = 0$. Since $\det(S) \neq 0$ we have $a \neq 0$ and $c \neq 0$. We use the notation of Lemma 5.2.1. We calculate the twisted Jacquet modules $(I^i/I^{i+1})_{N,\theta}$ for $i \in \{0,1,2\}$. Since the action of N on I^0/I^1 is trivial and θ is non-trivial, we have $(I^0/I^1)_{N,\theta} = 0$.
We consider the quotient $I^1/I^2 = \text{c-Ind}_H^P\sigma_1$, where

$$H = \left[\begin{smallmatrix}*&*&*&*\\&*&&*\\&&*&*\\&&&*\end{smallmatrix}\right],$$

and with σ_1 as in ii) of Lemma 5.2.1. We first show that for each function f in the standard model of this representation, the function $f^\circ : F \to V$, given by

$$f^\circ(w) = f(\left[\begin{smallmatrix}1&&&\\&1&w&\\&&1&\\&&&1\end{smallmatrix}\right]),$$

has compact support. Let K be a compact subset of P such that the support of f is contained in HK. If

$$\left[\begin{smallmatrix}1&&&\\&1&w&\\&&1&\\&&&1\end{smallmatrix}\right] = \left[\begin{smallmatrix}t&*&y&*\\&a&&*\\&&d&*\\&&&adt^{-1}\end{smallmatrix}\right]\left[\begin{smallmatrix}k_1&k_2&x_1&x_2\\k_3&k_4&x_3&x_4\\&&k_5&k_6\\&&k_7&k_8\end{smallmatrix}\right],$$

with the rightmost matrix being in K, then calculations show that $k_3 = k_7 = 0$ and $w = k_4^{-1}x_3$. Since k_4^{-1} and x_3 vary in bounded subsets, w is confined to a compact subset of F. This proves our assertion that f° has compact support. Next, for each function f in the standard model of c-$\text{Ind}_H^P\sigma_1$, consider the function $\tilde{f} : F^2 \to V$ given by

$$\tilde{f}(u,w) = f(\left[\begin{smallmatrix}1&&&\\u&1&w&\\&&1&\\&&-u&1\end{smallmatrix}\right])$$

for u,w in F. Let W be the space of all such functions $\tilde{f}$. Since the map $f \mapsto \tilde{f}$ is injective, we get a vector space isomorphism c-$\text{Ind}_H^P\sigma_1 \cong W$. In this new model, the action of N is given by

$$(\left[\begin{smallmatrix}1&&y&z\\&1&x&y\\&&1&\\&&&1\end{smallmatrix}\right]\tilde{f})(u,w) = \pi(\left[\begin{smallmatrix}1&y+uz\\&1\end{smallmatrix}\right])\tilde{f}(u,w+x+2uy+u^2z) \tag{98}$$

for x, y, z, u, w in F.
We claim that W contains $\mathcal{S}(F^2, V)$. Since W is translation invariant, it is enough to prove that W contains the function

$$f_{N,v}(u,w) = \begin{cases} v & \text{if } u, w \in \mathfrak{p}^N, \\ 0 & \text{if } u \notin \mathfrak{p}^N \text{ or } w \notin \mathfrak{p}^N, \end{cases}$$

for any v in V and any positive integer N. Again by translation invariance, we may assume that N is large enough so that

$$\sigma_1(h)v = v \qquad \text{for } h \in H \cap \Gamma_N, \tag{99}$$

where

$$\Gamma_N = \begin{bmatrix} 1+\mathfrak{p}^N & \mathfrak{p}^N & \mathfrak{p}^N & \mathfrak{p}^N \\ \mathfrak{p}^N & 1+\mathfrak{p}^N & \mathfrak{p}^N & \mathfrak{p}^N \\ & & 1+\mathfrak{p}^N & \mathfrak{p}^N \\ & & \mathfrak{p}^N & 1+\mathfrak{p}^N \end{bmatrix} \cap P. \tag{100}$$

Define $f:\ P \to V$ by

$$f(g) = \begin{cases} \sigma_1(h)v & \text{if } g = hk \text{ with } h \in H,\ k \in \Gamma_N, \\ 0 & g \notin H\Gamma_N. \end{cases}$$

Then, by (99), f is a well-defined element of c-$\mathrm{Ind}_H^Q \sigma_1$. It is easy to verify that $\tilde{f} = f_{N,v}$. This proves our claim that W contains $\mathcal{S}(F^2, V)$.
Now consider the map

$$W \longrightarrow \mathcal{S}(F, V), \qquad \tilde{f} \longmapsto \Big(w \mapsto f(\begin{bmatrix} 1 & & & \\ & 1 & w & \\ & & 1 & \\ & & & 1 \end{bmatrix} s_1)\Big), \tag{101}$$

where s_1 is defined in (7). This map is well-defined, since the function on the right is $(s_1 f)^\circ$, which we showed above has compact support. Similar considerations as above show that the map (101) is surjective.
We claim that the kernel of (101) is $\mathcal{S}(F^2, V)$. First suppose that $\tilde{f}$ lies in the kernel; we have to show that $\tilde{f}$ has compact support. Choose N large enough so that f is right invariant under Γ_N. Then, for u not in $\mathfrak{p}^{-N}$ and w in F,

$$\begin{aligned}
\tilde{f}(u,w) &= f(\begin{bmatrix} 1 & & & \\ u & 1 & w & \\ & & 1 & \\ & & -u & 1 \end{bmatrix}) \\
&= f(\begin{bmatrix} 1 & & & \\ & 1 & w & \\ & & 1 & \\ & & & 1 \end{bmatrix} \begin{bmatrix} 1 & u^{-1} & & \\ & 1 & & \\ & & 1 & -u^{-1} \\ & & & 1 \end{bmatrix} \begin{bmatrix} -u^{-1} & & & \\ & u & & \\ & & u^{-1} & \\ & & & -u \end{bmatrix} s_1 \begin{bmatrix} 1 & u^{-1} & & \\ & 1 & & \\ & & 1 & -u^{-1} \\ & & & 1 \end{bmatrix}) \\
&= f(\begin{bmatrix} 1 & u^{-1} & & \\ & 1 & & \\ & & 1 & -u^{-1} \\ & & & 1 \end{bmatrix} \begin{bmatrix} 1 & & -u^{-1}w & u^{-2}w \\ & 1 & w & -u^{-1}w \\ & & 1 & \\ & & & 1 \end{bmatrix} \begin{bmatrix} -u^{-1} & & & \\ & u & & \\ & & u^{-1} & \\ & & & -u \end{bmatrix} s_1) \\
&= \pi(\begin{bmatrix} 1 & -u^{-1}w \\ & 1 \end{bmatrix}) f(\begin{bmatrix} 1 & & & \\ & 1 & w & \\ & & 1 & \\ & & & 1 \end{bmatrix} \begin{bmatrix} -u^{-1} & & & \\ & u & & \\ & & u^{-1} & \\ & & & -u \end{bmatrix} s_1)
\end{aligned}$$

$$
\begin{aligned}
&= \pi(\begin{bmatrix} 1 & w^2u^{-1} \\ & 1 \end{bmatrix}) f(\begin{bmatrix} 1 & -w & & \\ & 1 & & \\ & & 1 & w \\ & & & 1 \end{bmatrix} \begin{bmatrix} 1 & & & \\ & -u^{-1} & & \\ & & -u & \\ & & & 1 \end{bmatrix} s_2) \\
&= \pi(\begin{bmatrix} 1 & -u^{-1}w \\ & 1 \end{bmatrix}) f(\begin{bmatrix} -u^{-1} & & & \\ & u & & \\ & & u^{-1} & \\ & & & -u \end{bmatrix} \begin{bmatrix} 1 & & & \\ & 1 & u^{-2}w & \\ & & 1 & \\ & & & 1 \end{bmatrix} s_1) \\
&= |u|^{-3} \pi(\begin{bmatrix} -u^{-1} & -u^{-2}w \\ & u^{-1} \end{bmatrix}) f(\begin{bmatrix} 1 & & & \\ & 1 & u^{-2}w & \\ & & 1 & \\ & & & 1 \end{bmatrix} s_1).
\end{aligned}
$$

This last expression is zero by assumption. For fixed u in $\mathfrak{p}^{-N}$, the function $\tilde{f}(u, \cdot)$ has compact support; this follows because each f° has compact support. Combining these facts shows that $\tilde{f}$ has compact support. Conversely, assume $\tilde{f}$ is in $\mathcal{S}(F^2, V)$. Then we can find a large enough N such that if u has valuation $-N$, the function $\tilde{f}(u, \cdot)$ is zero. Looking at the above calculation, we see that, for fixed such u,

$$
f(\begin{bmatrix} 1 & & & \\ & 1 & u^{-2}w & \\ & & 1 & \\ & & & 1 \end{bmatrix} s_1) = 0
$$

for all w in F. This shows that $\tilde{f}$ is in the kernel of the map (101), completing the proof of our claim about this kernel. We now have an exact sequence

$$
0 \longrightarrow \mathcal{S}(F^2, V) \longrightarrow W \longrightarrow \mathcal{S}(F, V) \longrightarrow 0. \tag{102}
$$

Note that the space $\mathcal{S}(F^2, V)$ is invariant under the action (98) of N. A calculation shows that the action of N on $\mathcal{S}(F, V)$ is given by

$$
(\begin{bmatrix} 1 & & y & z \\ & 1 & x & y \\ & & 1 & \\ & & & 1 \end{bmatrix} f)(w) = \pi(\begin{bmatrix} 1 & y \\ & 1 \end{bmatrix}) f(w + z) \tag{103}
$$

for x, y, z, w in F and f in $\mathcal{S}(F, V)$. Since the action of x is trivial and $a \neq 0$, it follows that $\mathcal{S}(F, V)_{N,\theta} = 0$. Hence, by (102), we have $W_{N,\theta} \cong \mathcal{S}(F^2, V)_{N,\theta}$. We will compute the Jacquet module $\mathcal{S}(F^2, V)_{N,\theta}$ in stages. The action of N on $\mathcal{S}(F^2, V)$ is given by (98). By ii) of Lemma 5.1.1, the map $\tilde{f} \mapsto f'$, where $f' : F \to V$ is given by

$$
f'(u) = \int\limits_F \psi^a(-u) \tilde{f}(u, w) \, dw,
$$

defines a vector space isomorphism

$$
W_{\begin{bmatrix} 1 & & & \\ & 1 & * & \\ & & 1 & \\ & & & 1 \end{bmatrix}, \psi^a} \xrightarrow{\sim} \mathcal{S}(F, V).
$$

The transfer of the action of the remaining group $\begin{bmatrix} 1 & & * & * \\ & 1 & & * \\ & & 1 & \\ & & & 1 \end{bmatrix}$ to $\mathcal{S}(F, V)$ is given by

$$
(\begin{bmatrix} 1 & & y & z \\ & 1 & & y \\ & & 1 & \\ & & & 1 \end{bmatrix} f)(u) = \psi(a(2uy + u^2 z)) \pi(\begin{bmatrix} 1 & y+uz \\ & 1 \end{bmatrix}) f(u)
$$

for $u, y, z \in F$ and $f \in \mathcal{S}(F, V)$. The subspace $\mathcal{S}(F^\times, V)$ of elements f of $\mathcal{S}(F, V)$ such that $f(0) = 0$ is preserved under this action, so that we have an exact sequence

$$0 \longrightarrow \mathcal{S}(F^\times, V) \longrightarrow \mathcal{S}(F, V) \longrightarrow \mathcal{S}(F, V)/\mathcal{S}(F^\times, V) \longrightarrow 0 \tag{104}$$

of representations of the group $\begin{bmatrix} 1 & & * & * \\ & 1 & & * \\ & & 1 & \\ & & & 1 \end{bmatrix}$. There is an isomorphism of vector spaces $\mathcal{S}(F, V)/\mathcal{S}(F^\times, V) \xrightarrow{\sim} V$ that sends f to $f(0)$. The transfer of the action of the group $\begin{bmatrix} 1 & & * & * \\ & 1 & & * \\ & & 1 & \\ & & & 1 \end{bmatrix}$ to V is given by

$$\begin{bmatrix} 1 & & y & z \\ & 1 & & y \\ & & 1 & \\ & & & 1 \end{bmatrix} v = \pi(\left[\begin{smallmatrix} 1 & y \\ & 1 \end{smallmatrix}\right])v \tag{105}$$

for $y, z \in F$ and $v \in V$. The non-vanishing of c and (105) imply that

$$(\mathcal{S}(F, V)/\mathcal{S}(F^\times, V))_{\left[\begin{smallmatrix} 1 & & & * \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{smallmatrix}\right], \psi^c} = 0.$$

Therefore,

$$\big(\mathcal{S}(F, V)/\mathcal{S}(F^\times, V)\big)_{\left[\begin{smallmatrix} 1 & & * & * \\ & 1 & & * \\ & & 1 & \\ & & & 1 \end{smallmatrix}\right], \theta} = 0.$$

Next, we define a vector space isomorphism of $\mathcal{S}(F^\times, V)$ with itself and then transfer the action. For f in $\mathcal{S}(F^\times, V)$, set $\tilde{f}(u) = \pi(\left[\begin{smallmatrix} u & \\ & 1 \end{smallmatrix}\right])f(u)$ for u in $F^\times$. The map defined by $f \mapsto \tilde{f}$ is an automorphism of vector spaces. The transfer of the action of $\begin{bmatrix} 1 & & * & * \\ & 1 & & * \\ & & 1 & \\ & & & 1 \end{bmatrix}$ is given by

$$(\begin{bmatrix} 1 & & y & z \\ & 1 & & y \\ & & 1 & \\ & & & 1 \end{bmatrix} f)(u) = \psi(a(2uy + u^2 z))\pi(\left[\begin{smallmatrix} 1 & uy+u^2z \\ & 1 \end{smallmatrix}\right])f(u)$$

for $f \in \mathcal{S}(F^\times, V)$, $y, z \in F$, and $u \in F^\times$. Now define a linear map

$$p : \mathcal{S}(F^\times, V) \longrightarrow \mathcal{S}(F^\times, V_{\left[\begin{smallmatrix} 1 & * \\ & 1 \end{smallmatrix}\right], \psi^{-2a}})$$

by composing the elements of $\mathcal{S}(F^\times, V)$ with the natural projection from V to $V_{\left[\begin{smallmatrix} 1 & * \\ & 1 \end{smallmatrix}\right], \psi^{-2a}} = V/V(\left[\begin{smallmatrix} 1 & * \\ & 1 \end{smallmatrix}\right], \psi^{-2a})$. The map p is surjective. Let f be in $\mathcal{S}(F^\times, V)$. Since f has compact support and is locally constant, we see that f is in the kernel of p if and only if

$$\text{there exists } l > 0 \text{ such that } \int\limits_{\mathfrak{p}^{-l}} \psi(2ay)\pi(\left[\begin{smallmatrix} 1 & y \\ & 1 \end{smallmatrix}\right])f(u)\, dy = 0 \qquad \text{for all } u \in F^\times. \tag{106}$$

Also, f is in $\mathcal{S}(F^\times, V)(\begin{bmatrix} 1 & & * & \\ & 1 & & * \\ & & 1 & \\ & & & 1 \end{bmatrix})$ if and only if

$$\text{there exists } k > 0 \text{ such that } \int\limits_{\mathfrak{p}^{-k}} \psi(2auy)\pi(\left[\begin{smallmatrix} 1 & uy \\ & 1 \end{smallmatrix}\right])f(u)\,dy = 0 \quad \text{for all } u \in F^\times. \tag{107}$$

Since f is locally constant and compactly supported the conditions (106) and (107) are equivalent. It follows that p induces an isomorphism of vector spaces:

$$\mathcal{S}(F^\times, V)_{\begin{bmatrix} 1 & & * & \\ & 1 & & * \\ & & 1 & \\ & & & 1 \end{bmatrix}, \psi^b} = \mathcal{S}(F^\times, V)_{\begin{bmatrix} 1 & & * & \\ & 1 & & * \\ & & 1 & \\ & & & 1 \end{bmatrix}} \xrightarrow{\sim} \mathcal{S}(F^\times, V_{\left[\begin{smallmatrix} 1 & * \\ & 1 \end{smallmatrix}\right], \psi^{-2a}}).$$

Transferring the action of $\begin{bmatrix} 1 & & & * \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix}$ on the first space to the last space results in the formula

$$\begin{bmatrix} 1 & & & z \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} f(u) = \psi(au^2z)\pi(\left[\begin{smallmatrix} 1 & u^2z \\ & 1 \end{smallmatrix}\right])f(u) = \psi(-au^2z)f(u), \qquad z \in F,\ u \in F^\times,$$

for f in $\mathcal{S}(F^\times, V_{\left[\begin{smallmatrix} 1 & * \\ & 1 \end{smallmatrix}\right], \psi^{-2a}})$.

Assume that L is a field; we will prove that

$$\mathcal{S}(F^\times, V_{\left[\begin{smallmatrix} 1 & * \\ & 1 \end{smallmatrix}\right], \psi^{-2a}})_{\begin{bmatrix} 1 & & & * \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix}, \psi^c} = 0. \tag{108}$$

Let f be in $\mathcal{S}(F^\times, V_{\left[\begin{smallmatrix} 1 & * \\ & 1 \end{smallmatrix}\right], \psi^{-2a}})$. Since the support of f is compact, and since there exists no u in $F^\times$ such that $c + au^2 = 0$ as $D = b^2/4 - ac = -ac$ is not in $F^{\times 2}$, there exists a positive integer l such that

$$\int\limits_{\mathfrak{p}^{-l}} \psi(-(c + au^2)z)\,dz = 0 \tag{109}$$

for u in the support of f. Hence, for u in $F^\times$,

$$\Big(\int\limits_{\mathfrak{p}^{-l}} \psi(-cz) \begin{bmatrix} 1 & & & z \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} f\,dz\Big)(u) = \Big(\int\limits_{\mathfrak{p}^{-l}} \psi(-(c + au^2)z)\,dz\Big)f(u) = 0. \tag{110}$$

This proves (108), and completes the argument that $(I^1/I^2)_{N,\theta} = 0$ in the case L is a field.

Now assume that L is not a field. We may further assume that $a = 1$ and $c = -1$ while retaining $b = 0$. The group $T = T_{\left[\begin{smallmatrix} a & b/2 \\ b/2 & c \end{smallmatrix}\right]} = T_{\left[\begin{smallmatrix} 1 & \\ & -1 \end{smallmatrix}\right]}$ consists of the elements

$$t = \begin{bmatrix} x & y & & \\ y & x & & \\ & & x & -y \\ & & -y & x \end{bmatrix} \tag{111}$$

with $x, y \in F$ such that $x^2 \neq y^2$. Define

$$\mathcal{S}(F^\times, V_{\left[\begin{smallmatrix}1&*\\&1\end{smallmatrix}\right],\psi^{-2a}}) \longrightarrow V_{\left[\begin{smallmatrix}1&*\\&1\end{smallmatrix}\right],\psi^{-2a}} \oplus V_{\left[\begin{smallmatrix}1&*\\&1\end{smallmatrix}\right],\psi^{-2a}} \tag{112}$$

by $f \mapsto f(1) \oplus f(-1)$. We assert that the kernel of this linear map is

$$\mathcal{S}(F^\times, V_{\left[\begin{smallmatrix}1&*\\&1\end{smallmatrix}\right],\psi^{-2a}})(\begin{bmatrix}1&&&*\\&1&&\\&&1&\\&&&1\end{bmatrix}, \psi^c).$$

Evidently, this subspace is contained in the kernel. Conversely, let $f \in \mathcal{S}(F^\times, V_{\left[\begin{smallmatrix}1&*\\&1\end{smallmatrix}\right],\psi^{-2a}})$ be such that $f(1) = f(-1) = 0$. Then there exists a positive integer l such that (109) holds for u in the support of f, implying that (110) holds. This proves our assertion. The map (112) is clearly surjective, so that we obtain an isomorphism

$$\mathcal{S}(F^\times, V_{\left[\begin{smallmatrix}1&*\\&1\end{smallmatrix}\right],\psi^{-2a}})_{\begin{bmatrix}1&&&*\\&1&&\\&&1&\\&&&1\end{bmatrix},\psi^c} \xrightarrow{\sim} V_{\left[\begin{smallmatrix}1&*\\&1\end{smallmatrix}\right],\psi^{-2a}} \oplus V_{\left[\begin{smallmatrix}1&*\\&1\end{smallmatrix}\right],\psi^{-2a}}.$$

We now have an isomorphism $(I^1/I^2)_{N,\theta} \xrightarrow{\sim} V_{\left[\begin{smallmatrix}1&*\\&1\end{smallmatrix}\right],\psi^{-2a}} \oplus V_{\left[\begin{smallmatrix}1&*\\&1\end{smallmatrix}\right],\psi^{-2a}}$. A calculation shows that the transfer of the action of T to $V_{\left[\begin{smallmatrix}1&*\\&1\end{smallmatrix}\right],\psi^{-2a}} \oplus V_{\left[\begin{smallmatrix}1&*\\&1\end{smallmatrix}\right],\psi^{-2a}}$ is given by

$$t(v_1 \oplus v_2) = \Big|\frac{x-y}{x+y}\Big|^{1/2} \sigma\big((x-y)(x+y)\big)\omega_\pi(x-y)v_1 \oplus$$
$$\oplus \Big|\frac{x-y}{x+y}\Big|^{-1/2} \sigma\big((x-y)(x+y)\big)\omega_\pi(x+y)v_2$$

for t as in (111) and $v_1, v_2 \in V_{\left[\begin{smallmatrix}1&*\\&1\end{smallmatrix}\right],\psi^{-2a}}$. Finally, the result stated in ii) is written with respect to $S = \left[\begin{smallmatrix}&1/2\\1/2&\end{smallmatrix}\right]$. To change to this choice note that the map

$$C : (I^1/I^2)_{N,\theta_{\left[\begin{smallmatrix}a&b/2\\b/2&c\end{smallmatrix}\right]}} \longrightarrow (I^1/I^2)_{N,\theta_{\left[\begin{smallmatrix}&1/2\\1/2&\end{smallmatrix}\right]}}$$

defined by $v \mapsto \left[\begin{smallmatrix}g&\\&g'\end{smallmatrix}\right] v$, where $g = \left[\begin{smallmatrix}-1&1\\1&1\end{smallmatrix}\right]$, is a well-defined isomorphism; recall that $a = 1, b = 0, c = -1$. Moreover, $C(tv) = t'C(v)$ for t as in (111) and

$$t' = \begin{bmatrix}x-y&&&\\&x+y&&\\&&x-y&\\&&&x+y\end{bmatrix} \in T_{\left[\begin{smallmatrix}&1/2\\1/2&\end{smallmatrix}\right]}.$$

It follows that the group $T_{\left[\begin{smallmatrix}&1/2\\1/2&\end{smallmatrix}\right]}$ acts on the isomorphic vector spaces

$$(I^1/I^2)_{N,\theta_{\left[\begin{smallmatrix}&1/2\\1/2&\end{smallmatrix}\right]}} \cong V_{\left[\begin{smallmatrix}1&*\\&1\end{smallmatrix}\right],\psi^{-2a}} \oplus V_{\left[\begin{smallmatrix}1&*\\&1\end{smallmatrix}\right],\psi^{-2a}} \cong V_{\left[\begin{smallmatrix}1&*\\&1\end{smallmatrix}\right],\psi} \oplus V_{\left[\begin{smallmatrix}1&*\\&1\end{smallmatrix}\right],\psi}$$

via the formula in ii).

Next, we consider the quotient $I^2/I^3 = \text{c-Ind}_M^P(\sigma_2)$ from iii) of Lemma 5.2.1. By Lemma 5.1.2, restriction of functions in the standard model of this representation to N gives an N-isomorphism $\text{c-Ind}_M^P(\sigma_2) \cong \mathcal{S}(N,V)$. An application of i) and ii) of Lemma 5.1.1 shows that $\mathcal{S}(N,V)_{N,\theta} \cong V$ via the map defined by

$$f \longmapsto \int\limits_N \theta(n)^{-1} f(n)\, dn.$$

Transferring the action of T we find that $t \in T$ acts by $\sigma_2(t)$ on V. If $t = \begin{bmatrix} g & \\ & \det(g)g' \end{bmatrix}$ as in (16), then

$$\sigma_2(t) = \sigma(\det(g))\omega_\pi(\det(g))\pi(g').$$

This concludes the proof. ∎

In case of a one-dimensional representation of M, it follows from this lemma that

$$(\chi 1_{\mathrm{GL}(2)} \rtimes \sigma)_{N,\theta} = \mathbb{C}_{(\sigma\chi)\circ \mathrm{N}_{L/F}} \tag{113}$$

as T-modules. In case L is a field and π is irreducible, it follows from Lemma 5.2.2 that

$$\mathrm{Hom}_T((\pi \rtimes \sigma)_{N,\theta}, \mathbb{C}_\Lambda) = \mathrm{Hom}_T(\sigma\pi, \mathbb{C}_\Lambda). \tag{114}$$

Hence, in view of (20), the space of (Λ,θ)-Bessel functionals on $\pi \rtimes \sigma$ is isomorphic to the space of (Λ,θ)-Waldspurger functionals on $\sigma\pi$.

5.3 Klingen induced representations

Let π be an admissible representation of $\mathrm{GL}(2,F)$, let χ be a character of $F^\times$, and let $\chi \rtimes \pi$ be as defined in Sect. 1.2; see (10). In this section we will calculate the twisted Jacquet modules $(\chi \rtimes \pi)_{N,\theta}$ for any non-degenerate character θ of N as a module of T. In the split case our results make several corrections to Proposition 2.3 and Proposition 2.4 of [23].

5.3.1 Lemma. *Let χ be a character of $F^\times$ and π an admissible representation of $\mathrm{GL}(2,F)$. Let I be the space of the Klingen induced representation $\chi \rtimes \pi$. There is a filtration of P-modules*

$$I^2 = 0 \subset I^1 \subset I^0 = I.$$

with the quotients given as follows.

i) $I^0/I^1 = \text{c-Ind}_B^P \sigma_0$, *where*

$$\sigma_0(\begin{bmatrix} t & * & * & * \\ & a & b & * \\ & & d & * \\ & & & adt^{-1} \end{bmatrix}) = \chi(t)\,|t|^2\,|ad|^{-1}\,\pi(\begin{bmatrix} a & b \\ & d \end{bmatrix})$$

for b in F and a,d,t in $F^\times$.

ii) $I^1/I^2 = \text{c-Ind}^P_{\left[\begin{smallmatrix} * & * & & * \\ & * & & \\ & & * & * \\ & & & * \end{smallmatrix}\right]} \sigma_1$, *where*

$$\sigma_1(\begin{bmatrix} t & * & & x \\ & a & & \\ & & d & * \\ & & & adt^{-1} \end{bmatrix}) = \chi(d)\,|a^{-1}d|\,\pi(\left[\begin{smallmatrix} t & x \\ & adt^{-1} \end{smallmatrix}\right])$$

for x in F and a, d, t in $F^\times$.

Proof. This follows by going through the procedure of Sections 6.2 and 6.3 of [5]. ∎

5.3.2 LEMMA. *Let χ be a character of $F^\times$, and let (π, V) be an admissible representation of $\mathrm{GL}(2, F)$. We assume that π has a central character ω_π. Let I be the standard space of the Klingen induced representation $\chi \rtimes \pi$. Let N be the unipotent radical of the Siegel parabolic subgroup, and let θ be the character of N defined in (15). We assume that the associated quadratic extension L is a field. Then, as T-modules,*

$$I_{N,\theta} \cong \bigoplus_{\Lambda|_{F^\times} = \chi\omega_\pi} d \cdot \Lambda, \qquad \text{where } d = \dim \mathrm{Hom}_{\left[\begin{smallmatrix} 1 & * \\ & 1 \end{smallmatrix}\right]}(\pi, \psi).$$

In particular, $I_{N,\theta} = 0$ if π is one-dimensional.

Proof. We will first prove that $(I_0/I_1)_{N,\theta} = 0$, where the notations are as in Lemma 5.3.1. We may assume that the element b appearing in the matrix S in (15) is zero. For f in the standard space of the induced representation $I_0/I_1 = \text{c-Ind}^P_B \sigma_0$, let

$$\tilde{f}(u) = f(\begin{bmatrix} 1 & & & \\ u & 1 & & \\ & & 1 & \\ & & -u & 1 \end{bmatrix}), \qquad u \in F.$$

Let W be the space of all functions $F \to \mathbb{C}$ of the form $\tilde{f}$, where f runs through $\text{c-Ind}^P_B \sigma_0$. Since the map $f \mapsto \tilde{f}$ is injective, we obtain a vector space isomorphism $\text{c-Ind}^P_B \sigma_0 \cong W$. The identity

$$\begin{bmatrix} 1 & & & \\ u & 1 & & \\ & & 1 & \\ & & -u & 1 \end{bmatrix} = \begin{bmatrix} 1 & u^{-1} & & \\ & 1 & & \\ & & 1 & -u^{-1} \\ & & & 1 \end{bmatrix} \begin{bmatrix} -u^{-1} & & & \\ & u & & \\ & & u^{-1} & \\ & & & -u \end{bmatrix} s_1 \begin{bmatrix} 1 & u^{-1} & & \\ & 1 & & \\ & & 1 & -u^{-1} \\ & & & 1 \end{bmatrix},$$

where s_1 is as in (7), shows that $\tilde{f}$ satisfies

$$\tilde{f}(u) = \chi(-u^{-1})|u|^{-2}\pi(\left[\begin{smallmatrix} u & \\ & u^{-1} \end{smallmatrix}\right])f(s_1) \qquad \text{for } |u| \gg 0. \tag{115}$$

The space W consists of locally constant functions. Furthermore, W is invariant under translations, i.e., if $f' \in W$, then the function $u \mapsto f'(u + x)$ is also in W, for any x in F.

We claim that W contains $\mathcal{S}(F,V)$. Since W is translation invariant, it is enough to prove that W contains the function

$$f_{N,v}(u)=\begin{cases} v & \text{if } u\in\mathfrak{p}^N,\\ 0 & \text{if } u\notin\mathfrak{p}^N,\end{cases}$$

for any v in V and any positive integer N. Again by translation invariance, we may assume that N is large enough so that

$$\sigma_0(b)v=v \qquad \text{for } b\in B\cap\Gamma_N, \tag{116}$$

where Γ_N is as in (100). Define $f:\,P\to V$ by

$$f(g)=\begin{cases}\sigma_0(b)v & \text{if } g=bk \text{ with } b\in B,\ k\in\Gamma_N,\\ 0 & g\notin B\Gamma_N.\end{cases}$$

Then, by (116), f is a well-defined element of c-$\mathrm{Ind}_B^P\sigma_0$. It is easy to verify that $\tilde{f}=f_{N,v}$. This proves our claim that W contains $\mathcal{S}(F,V)$.
We define a linear map $W\to V$ by sending $\tilde{f}$ to the vector $f(s_1)$ in (115). Then the kernel of this map is $\mathcal{S}(F,V)$. We claim that the map is surjective. To see this, let v be in V. Again choose N large enough so that (116) holds. Then the function $f:\,P\to V$ given by

$$f(g)=\begin{cases}\sigma_0(b)v & \text{if } g=bs_1k \text{ with } b\in B,\ k\in\Gamma_N,\\ 0 & g\notin Bs_1\Gamma_N.\end{cases}$$

is a well-defined element of c-$\mathrm{Ind}_B^P\sigma_0$ with $f(s_1)=v$. This proves our claim that the map $W\to V$ is surjective. We therefore get an exact sequence

$$0\longrightarrow\mathcal{S}(F,V)\longrightarrow W\longrightarrow V\longrightarrow 0. \tag{117}$$

The transfer of the action of N to W is given by

$$\Big(\begin{bmatrix}1&&y&z\\&1&x&y\\&&1&\\&&&1\end{bmatrix}\tilde{f}\Big)(u)=\pi\Big(\begin{bmatrix}1&x+2uy+u^2z\\&1\end{bmatrix}\Big)\tilde{f}(u)$$

for all x,y,z,u in F. Evidently, the subspace $\mathcal{S}(F,V)$ is invariant under N. Moreover, the action of N on V is given by

$$\begin{bmatrix}1&&y&z\\&1&x&y\\&&1&\\&&&1\end{bmatrix}v=\pi\Big(\begin{bmatrix}1&z\\&1\end{bmatrix}\Big)v \tag{118}$$

for all x,y,z in F and v in V.
To prove that $(I_0/I_1)_{N,\theta}=0$, it suffices to show that $\mathcal{S}(F,V)_{N,\theta}=0$ and $V_{N,\theta}=0$. Since the element a in the matrix S is non-zero, it follows from (118) that $V_{N,\theta}=0$.

To prove that $\mathcal{S}(F,V)_{N,\theta}=0$, we define a map p from $\mathcal{S}(F,V)$ to

$$\mathcal{S}(F,V_{\left[\begin{smallmatrix}1&*\\&1\end{smallmatrix}\right],\psi^a})=\mathcal{S}(F,V/V(\left[\begin{smallmatrix}1&*\\&1\end{smallmatrix}\right],\psi^a))$$

by sending f to f composed with the projection from V to $V/V(\left[\begin{smallmatrix}1&*\\&1\end{smallmatrix}\right],\psi^a)$. This map is surjective. It is easy to see that p induces an isomorphism

$$\mathcal{S}(F,V)_{\left[\begin{smallmatrix}1&&&\\&1&*&\\&&1&\\&&&1\end{smallmatrix}\right],\psi^a}\cong\mathcal{S}(F,V_{\left[\begin{smallmatrix}1&*\\&1\end{smallmatrix}\right],\psi^a}).$$

For the space on the right we have the action

$$(\left[\begin{smallmatrix}1&&y&z\\&1&&y\\&&1&\\&&&1\end{smallmatrix}\right]f)(u)=\pi(\left[\begin{smallmatrix}1&2uy+u^2z\\&1\end{smallmatrix}\right])f(u),\qquad u\in F.$$

By iii) of Lemma 5.1.1, the map $f\mapsto f(0)$ induces an isomorphism

$$\mathcal{S}(F,V_{\left[\begin{smallmatrix}1&*\\&1\end{smallmatrix}\right],\psi^a})_{\left[\begin{smallmatrix}1&&*&\\&1&&*\\&&1&\\&&&1\end{smallmatrix}\right]}\cong V_{\left[\begin{smallmatrix}1&*\\&1\end{smallmatrix}\right],\psi^a}.$$

For the space on the right we have the action $\left[\begin{smallmatrix}1&&&z\\&1&&\\&&1&\\&&&1\end{smallmatrix}\right]v=v$. Taking a twisted Jacquet module with respect to the character ψ^c gives zero, since $c\neq0$. This concludes our proof that $(I_0/I_1)_{N,\theta}=0$.
Next let σ_1 be as in ii) of Lemma 5.3.1. Let

$$H_1=\left[\begin{smallmatrix}*&*&&*\\&*&&\\&&*&*\\&&&*\end{smallmatrix}\right]$$

and $H_2=TN$. By Lemma 2.3.1, we have $P=H_1H_2$. To verify the hypotheses of Lemma 5.1.2, let K be a compact subset of P. Write $P=MN$ and let $p:P\to N$ be the resulting projection map. Since p is continuous, the set $p(K)$ is compact. There exists a compact subset K_T of T such that $T=F^\times K_T$. Then $M\subset H_1K_T$ by Lemma 2.3.1. Therefore $K\subset H_1K_2$ with $K_2=K_Tp(K)$. By Lemma 5.1.2, restriction of functions gives a TN isomorphism

$$\text{c-Ind}^P_{\left[\begin{smallmatrix}*&*&&*\\&*&&\\&&*&*\\&&&*\end{smallmatrix}\right]}\sigma_1\cong\text{c-Ind}^{TN}_{F^\times Z^J}(\sigma_1\big|_{F^\times Z^J}).$$

Note that $F^\times$ acts via the character $\chi\omega_\pi$ on this module. Since T is compact modulo $F^\times$, the Jacquet module $(\text{c-Ind}^{TN}_{F^\times Z^J}(\sigma_1\big|_{F^\times Z^J}))_{N,\theta}$ is a direct sum over characters of T. Let Λ be a character of T. It is easy to verify that

$$\mathrm{Hom}_T\big((\text{c-Ind}^{TN}_{F^\times Z^J}(\sigma_1\big|_{F^\times Z^J}))_{N,\theta},\Lambda\big)=\mathrm{Hom}_{TN}\big(\text{c-Ind}^{TN}_{F^\times Z^J}(\sigma_1\big|_{F^\times Z^J}),\Lambda\otimes\theta\big).$$

By Frobenius reciprocity, the space on the right is isomorphic to

$$\mathrm{Hom}_{F^\times Z^J}\big(\sigma_1\big|_{F^\times Z^J},(\Lambda\otimes\theta)\big|_{F^\times Z^J}\big).\tag{119}$$

This space is zero unless the restriction of Λ to $F^\times$ equals $\chi\omega_\pi$. Assume this is the case. Then (119) is equal to

$$\mathrm{Hom}_{\left[\begin{smallmatrix}1&*\\&1\end{smallmatrix}\right]}(\pi,\psi^c)\cong\mathrm{Hom}_{\left[\begin{smallmatrix}1&*\\&1\end{smallmatrix}\right]}(\pi,\psi).$$

This concludes the proof. ∎

5.3.3 Lemma. *Let χ be a character of $F^\times$ and π an admissible representation of $\mathrm{GL}(2,F)$. Let I be the space of the Klingen induced representation $\chi\rtimes\pi$. There is a filtration of Q-modules*

$$I^3=0\subset I^2\subset I^1\subset I^0=I.$$

with the quotients given as follows.

i) $I^0/I^1=\sigma_0$, where

$$\sigma_0(\begin{bmatrix}t&*&*&*\\&a&b&*\\&c&d&*\\&&&(ad-bc)t^{-1}\end{bmatrix})=\chi(t)\,|t^2(ad-bc)^{-1}|\,\pi(\left[\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right])$$

for $\left[\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right]$ in $\mathrm{GL}(2,F)$ and t in $F^\times$.

ii) $I^1/I^2=\text{c-Ind}^Q_{\left[\begin{smallmatrix}*&&*&*\\&*&*&*\\&&*&\\&&&*\end{smallmatrix}\right]}\sigma_1$, where

$$\sigma_1(\begin{bmatrix}t&&*&x\\&a&b&*\\&&d&\\&&&adt^{-1}\end{bmatrix})=\chi(a)\,|ad^{-1}|\,\pi(\left[\begin{smallmatrix}t&x\\&adt^{-1}\end{smallmatrix}\right])$$

for b,x in F and a,d,t in $F^\times$.

iii) $I^2/I^3=I^2=\text{c-Ind}^Q_{\left[\begin{smallmatrix}*&&&\\&*&*&\\&*&*&\\&&&*\end{smallmatrix}\right]}\sigma_2$, where

$$\sigma_2(\begin{bmatrix}t&&&\\&a&b&\\&c&d&\\&&&(ad-bc)t^{-1}\end{bmatrix})=\chi(t^{-1}(ad-bc))\,|t^{-2}(ad-bc)|\,\pi(\left[\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right])$$

for $\left[\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right]$ in $\mathrm{GL}(2,F)$ and t in $F^\times$.

Proof. This follows by going through the procedure of Sections 6.2 and 6.3 of [5]. ∎

5.3.4 Lemma. *Let χ be a character of $F^\times$, and let (π,V) be an admissible representation of $\mathrm{GL}(2,F)$. Let I be the standard space of the Klingen induced representation $\chi\rtimes\pi$. Let N be the unipotent radical of the Siegel parabolic subgroup, and let θ be the character of N defined in (23) (i.e., we consider the split case). Then there is a filtration*

$$0\subset J_3\subset J_2\subset J_1=I_{N,\theta},$$

with the quotients given as follows.

- $J_1/J_2 \cong V$
- $J_2/J_3 \cong V_{\left[\begin{smallmatrix}1&\\ *&1\end{smallmatrix}\right]}$.
- $J_3 \cong \mathcal{S}(F^\times, V_{\left[\begin{smallmatrix}1&\\ *&1\end{smallmatrix}\right],\psi})$.

The action of the stabilizer of θ is given as follows,

$$\begin{aligned}
\mathrm{diag}(a,b,a,b)v &= \chi(a)\pi(\left[\begin{smallmatrix}a&\\ &b\end{smallmatrix}\right])v && \text{for } v \in J_1/J_2,\\
\mathrm{diag}(a,b,a,b)v &= \chi(b)\pi(\left[\begin{smallmatrix}a&\\ &b\end{smallmatrix}\right])v && \text{for } v \in J_2/J_3,\\
(\mathrm{diag}(a,b,a,b)f)(u) &= \chi(b)\pi(\left[\begin{smallmatrix}a&\\ &a\end{smallmatrix}\right])f(a^{-1}bu) && \text{for } f \in J_3,\ u \in F^\times,
\end{aligned}$$

for all a and b in $F^\times$. In particular, we have the following special cases.

i) *Assume that $\pi = \sigma 1_{\mathrm{GL}(2)}$. Then the twisted Jacquet module $I_{N,\theta} = I/\langle \theta(n)v - \rho(n)v : n \in N,\ v \in I\rangle$ is two-dimensional. More precisely, there is a filtration*

$$0 \subset J_2 \subset J_1 = I_{N,\theta},$$

where J_2 and J_1/J_2 are both one-dimensional, and the action of the stabilizer of θ is given as follows,

$$\begin{aligned}
\mathrm{diag}(a,b,a,b)v &= \chi(a)\sigma(ab)v && \text{for } v \in J_1/J_2,\\
\mathrm{diag}(a,b,a,b)v &= \chi(b)\sigma(ab)v && \text{for } v \in J_2,
\end{aligned}$$

for all a and b in $F^\times$.

ii) *Assume that π is infinite-dimensional and irreducible. Then there is a filtration*

$$0 \subset J_3 \subset J_2 \subset J_1 = I_{N,\theta},$$

with the quotients given as follows.

- $J_1/J_2 \cong V$
- $J_2/J_3 \cong V_{\left[\begin{smallmatrix}1&\\ *&1\end{smallmatrix}\right]}$. *Hence, J_2/J_3 is 2-dimensional if π is a principal series representation, 1-dimensional if π is a twist of the Steinberg representation, and 0 if π is supercuspidal.*
- $J_3 \cong \mathcal{S}(F^\times)$.

The action of the stabilizer of θ is given as follows,

$$\begin{aligned}
\mathrm{diag}(a,b,a,b)v &= \chi(a)\pi(\left[\begin{smallmatrix}a&\\ &b\end{smallmatrix}\right])v && \text{for } v \in J_1/J_2,\\
\mathrm{diag}(a,b,a,b)v &= \chi(b)\pi(\left[\begin{smallmatrix}a&\\ &b\end{smallmatrix}\right])v && \text{for } v \in J_2/J_3,\\
(\mathrm{diag}(a,b,a,b)f)(u) &= \chi(b)\omega_\pi(a)f(a^{-1}bu) && \text{for } f \in J_3,\ u \in F^\times,
\end{aligned}$$

for all a and b in $F^\times$.

Proof. It will be easier to work with the conjugate subgroup $N_{\rm alt}$ and the character $\theta_{\rm alt}$ of $N_{\rm alt}$ defined in (26). For the top quotient from i) of Lemma 5.3.3 we have

$$(I^0/I^1)_{N_{\rm alt},\theta_{\rm alt}} = 0,$$

since the subgroup $\begin{bmatrix} 1 & * & & \\ & 1 & & \\ & & 1 & * \\ & & & 1 \end{bmatrix}$ acts trivially on I^0/I^1, but $\theta_{\rm alt}$ is not trivial on this subgroup. We consider the quotient $I^1/I^2 = \text{c-Ind}_H^Q\sigma_1$, where $H = \begin{bmatrix} * & & * & * \\ & * & * & * \\ & & * & \\ & & & * \end{bmatrix}$, and with σ_1 as in ii) of Lemma 5.3.3. We first show that for each function f in the standard model of this representation, the function $f^\circ : F \to V$, given by

$$f^\circ(w) = f(\begin{bmatrix} 1 & -w & & \\ & 1 & & \\ & & 1 & w \\ & & & 1 \end{bmatrix}),$$

has compact support. Let K be a compact subset of Q such that the support of f is contained in HK. If

$$\begin{bmatrix} 1 & -w & & \\ & 1 & & \\ & & 1 & w \\ & & & 1 \end{bmatrix} = \begin{bmatrix} t & & * & x \\ & a & b & * \\ & & d & \\ & & & adt^{-1} \end{bmatrix} \begin{bmatrix} k_0 & x_1 & x_2 & x_3 \\ & k_1 & k_2 & x_4 \\ & k_3 & k_4 & x_5 \\ & & & k_5 \end{bmatrix},$$

with the rightmost matrix being in K, then calculations show that $k_3 = 0$ and $w = k_4^{-1}x_5$. Since k_4^{-1} and x_5 vary in bounded subsets, w is confined to a compact subset of F. This proves our assertion that f° has compact support. Next, for each function f in the standard model of $\text{c-Ind}_H^Q\sigma_1$, consider the function $\tilde f : F^2 \to V$ given by

$$\tilde f(u,w) = f(\begin{bmatrix} 1 & -w & & \\ & 1 & & \\ & u & 1 & w \\ & & & 1 \end{bmatrix}).$$

Let W be the space of all such functions $\tilde f$. Since the map $f \mapsto \tilde f$ is injective, we get a vector space isomorphism $\text{c-Ind}_H^Q\sigma_1 \cong W$. Evidently, in this new model, the action of $N_{\rm alt}$ is given by

$$(\begin{bmatrix} 1 & -y & & z \\ & 1 & & \\ & x & 1 & y \\ & & & 1 \end{bmatrix} \tilde f)(u,w) = \tilde f(u+x, w+y). \tag{120}$$

We claim that W contains $\mathcal{S}(F^2,V)$. Since W is translation invariant, it is enough to prove that W contains the function

$$f_{N,v}(u,w) = \begin{cases} v & \text{if } u,w \in \mathfrak{p}^N, \\ 0 & \text{if } u \notin \mathfrak{p}^N \text{ or } w \notin \mathfrak{p}^N, \end{cases}$$

for any v in V and any positive integer N. Again by translation invariance, we may assume that N is large enough so that

$$\sigma_1(h)v = v \qquad \text{for } h \in H \cap \Gamma_N, \tag{121}$$

where

$$\Gamma_N = \begin{bmatrix} 1+\mathfrak{p}^N & \mathfrak{p}^N & \mathfrak{p}^N & \mathfrak{p}^N \\ & 1+\mathfrak{p}^N & \mathfrak{p}^N & \mathfrak{p}^N \\ & \mathfrak{p}^N & 1+\mathfrak{p}^N & \mathfrak{p}^N \\ & & & 1+\mathfrak{p}^N \end{bmatrix} \cap Q. \tag{122}$$

Define $f: Q \to V$ by

$$f(g) = \begin{cases} \sigma_1(h)v & \text{if } g = hk \text{ with } h \in H,\ k \in \Gamma_N, \\ 0 & g \notin H\Gamma_N. \end{cases}$$

Then, by (121), f is a well-defined element of c-$\mathrm{Ind}_H^Q \sigma_1$. It is easy to verify that $\tilde{f} = f_{N,v}$. This proves our claim that W contains $\mathcal{S}(F^2, V)$.
Now consider the map

$$W \longrightarrow \mathcal{S}(F, V), \qquad \tilde{f} \longmapsto \Big(w \mapsto f(\begin{bmatrix} 1 & -w & & \\ & 1 & & \\ & & 1 & w \\ & & & 1 \end{bmatrix} s_2)\Big), \tag{123}$$

where s_2 is defined in (7). This map is well-defined, since the function on the right is $(s_2 f)^\circ$, which we showed above has compact support. Similar considerations as above show that the map (123) is surjective.
We claim that the kernel of (123) is $\mathcal{S}(F^2, V)$. First suppose that $\tilde{f}$ lies in the kernel; we have to show that $\tilde{f}$ has compact support. Choose N large enough so that f is right invariant under Γ_N. Then, for u not in $\mathfrak{p}^{-N}$ and w in F,

$$\begin{aligned} \tilde{f}(u,w) &= f(\begin{bmatrix} 1 & -w & & \\ & 1 & & \\ & u & 1 & w \\ & & & 1 \end{bmatrix}) \\ &= f(\begin{bmatrix} 1 & -w & & \\ & 1 & & \\ & & 1 & w \\ & & & 1 \end{bmatrix} \begin{bmatrix} 1 & & & \\ & 1 & u^{-1} & \\ & & 1 & \\ & & & 1 \end{bmatrix} \begin{bmatrix} 1 & & & \\ & -u^{-1} & & \\ & & -u & \\ & & & 1 \end{bmatrix} s_2 \begin{bmatrix} 1 & & & \\ & 1 & u^{-1} & \\ & & 1 & \\ & & & 1 \end{bmatrix}) \\ &= f(\begin{bmatrix} 1 & & -wu^{-1} & w^2u^{-1} \\ & 1 & u^{-1} & -wu^{-1} \\ & & 1 & \\ & & & 1 \end{bmatrix} \begin{bmatrix} 1 & -w & & \\ & 1 & & \\ & & 1 & w \\ & & & 1 \end{bmatrix} \begin{bmatrix} 1 & & & \\ & -u^{-1} & & \\ & & -u & \\ & & & 1 \end{bmatrix} s_2) \\ &= \pi(\begin{bmatrix} 1 & w^2u^{-1} \\ & 1 \end{bmatrix}) f(\begin{bmatrix} 1 & -w & & \\ & 1 & & \\ & & 1 & w \\ & & & 1 \end{bmatrix} \begin{bmatrix} 1 & & & \\ & -u^{-1} & & \\ & & -u & \\ & & & 1 \end{bmatrix} s_2) \\ &= \chi(-u^{-1})|u|^{-2} \pi(\begin{bmatrix} 1 & w^2u^{-1} \\ & 1 \end{bmatrix}) f(\begin{bmatrix} 1 & wu^{-1} & & \\ & 1 & & \\ & & 1 & -wu^{-1} \\ & & & 1 \end{bmatrix} s_2). \end{aligned}$$

This last expression is zero by assumption. For fixed u in $\mathfrak{p}^{-N}$, the function $\tilde{f}(u, \cdot)$ has compact support; this follows because each f° has compact support. Combining these facts shows that $\tilde{f}$ has compact support. Conversely, assume $\tilde{f}$ is in $\mathcal{S}(F^2, V)$. Then we can find a large enough N such that if u has valuation $-N$, the function $\tilde{f}(u, \cdot)$ is zero. Looking at the above calculation, we see that, for fixed such u,

$$f(\begin{bmatrix} 1 & wu^{-1} & & \\ & 1 & & \\ & & 1 & -wu^{-1} \\ & & & 1 \end{bmatrix} s_2) = 0$$

for all w in F. This shows that $\tilde f$ is in the kernel of the map (123), completing the proof of our claim about this kernel. We now have an exact sequence

$$0 \longrightarrow \mathcal{S}(F^2,V) \longrightarrow W \longrightarrow \mathcal{S}(F,V) \longrightarrow 0. \tag{124}$$

Note that the space $\mathcal{S}(F^2,V)$ is invariant under the action (120) of $N_{\rm alt}$. A calculation shows that the action of $N_{\rm alt}$ on $\mathcal{S}(F,V)$ is given by

$$(\begin{bmatrix} 1&-y&&z\\ &1&&\\ &x&1&y\\ &&&1\end{bmatrix} f)(w) = \pi(\begin{bmatrix} 1&z-2wy-w^2x\\ &1\end{bmatrix})f(w) \tag{125}$$

for x,y,z,w in F and f in $\mathcal{S}(F,V)$.
We claim that $\mathcal{S}(F,V)_{N_{\rm alt},\theta_{\rm alt}}=0$. To prove this, we calculate this Jacquet module in stages. We define a map p from $\mathcal{S}(F,V)$ to

$$\mathcal{S}(F,V_{\left[\begin{smallmatrix}1&*\\&1\end{smallmatrix}\right]}) = \mathcal{S}(F,V/V([\begin{smallmatrix}1&*\\&1\end{smallmatrix}]))$$

by sending f to f composed with the natural projection from V to $V/V([\begin{smallmatrix}1&*\\&1\end{smallmatrix}])$. This map is surjective and has kernel $\mathcal{S}(F,V)(\begin{bmatrix}1&&&*\\&1&&\\&&1&\\&&&1\end{bmatrix})$. Hence, we obtain an isomorphism

$$\mathcal{S}(F,V)_{\left[\begin{smallmatrix}1&&&*\\&1&&\\&&1&\\&&&1\end{smallmatrix}\right]} \cong \mathcal{S}(F,V_{\left[\begin{smallmatrix}1&*\\&1\end{smallmatrix}\right]}).$$

The action of the group $\begin{bmatrix}1&*&&\\&1&&\\&*&1&*\\&&&1\end{bmatrix}$ on these spaces is trivial. Since $\theta_{\rm alt}$ is not trivial on this group, this proves our claim that $\mathcal{S}(F,V)_{N_{\rm alt},\theta_{\rm alt}}=0$.
By (124), we now have $W_{N_{\rm alt},\theta_{\rm alt}} \cong \mathcal{S}(F^2,V)_{N_{\rm alt},\theta_{\rm alt}}$. The action of $N_{\rm alt}$ on $\mathcal{S}(F^2,V)$ is given by (120). Since $\mathcal{S}(F^2,V)=\mathcal{S}(F)\otimes\mathcal{S}(F)\otimes V$, Lemma 5.1.1 implies that the map

$$f \longmapsto \int\limits_F \int\limits_F f(u,w)\psi(-w)\,du\,dw$$

induces an isomorphism $\mathcal{S}(F^2,V)_{N_{\rm alt},\theta_{\rm alt}} \cong V$. Moreover, a calculation shows that ${\rm diag}(a,a,b,b)$ acts on $\mathcal{S}(F^2,V)_{N_{\rm alt},\theta_{\rm alt}} \cong V$ by $\chi(a)\pi([\begin{smallmatrix}a&\\&b\end{smallmatrix}])$.
Finally, we consider the bottom quotient $I^2/I^3 = \text{c-Ind}^Q_{\left[\begin{smallmatrix}*&&&\\&*&*&\\&*&*&\\&&&*\end{smallmatrix}\right]}\sigma_2$ with σ_2 as in iii) of Lemma 5.3.3. If we associate with a function f in the standard model of this induced representation the function

$$\tilde f(u,v,w) = f(\begin{bmatrix}1&-v&u&w\\&1&&u\\&&1&v\\&&&1\end{bmatrix}),$$

then, by Lemma 5.1.2, we obtain an isomorphism $I^2/I^3 \cong \mathcal{S}(F^3,V)$. A calculation shows that the action of $N_{\rm alt}$ on $\mathcal{S}(F^3,V)$ is given by

$$(\begin{bmatrix} 1&-y&&z\\ &1&&\\ &x&1&y\\ &&&1\end{bmatrix} f)(u,v,w) = \pi([\begin{smallmatrix}1&\\x&1\end{smallmatrix}])f(u,v+y-ux,w+z+uy) \tag{126}$$

for x, y, z, u, v, w in F and f in $\mathcal{S}(F^3, V)$. This time we take Jacquet modules step by step, starting with the z-variable. Lemma 5.1.1 shows that the map

$$f \longmapsto \Big((u,v) \mapsto \int\limits_F f(u,v,w)\, dw \Big)$$

induces an isomorphism $\mathcal{S}(F^3,V)_{\left[\begin{smallmatrix} 1&&&*\\ &1&&\\ &&1&\\ &&&1 \end{smallmatrix}\right]} \cong \mathcal{S}(F^2, V)$. On $\mathcal{S}(F^2,V)$ we have the action

$$(\left[\begin{smallmatrix} 1&-y&&\\ &1&&\\ &x&1&y\\ &&&1 \end{smallmatrix}\right] f)(u,v) = \pi(\left[\begin{smallmatrix} 1&\\ x&1 \end{smallmatrix}\right]) f(u, v + y - ux)$$

for x, y, u, v in F and f in $\mathcal{S}(F^2, V)$. Part ii) of Lemma 5.1.1 shows that the map

$$f \longmapsto \Big(u \mapsto \int\limits_F f(u,v)\psi(-v)\, dv \Big)$$

induces an isomorphism $\mathcal{S}(F^2,V)_{\left[\begin{smallmatrix} 1&*&&\\ &1&&\\ &&1&*\\ &&&1 \end{smallmatrix}\right],\psi} \cong \mathcal{S}(F, V)$. A calculation shows that on $\mathcal{S}(F,V)$ we have the actions

$$(\left[\begin{smallmatrix} 1&&&\\ &1&&\\ &x&1&\\ &&&1 \end{smallmatrix}\right] f)(u) = \psi(-ux)\pi(\left[\begin{smallmatrix} 1&\\ x&1 \end{smallmatrix}\right]) f(u) \tag{127}$$

for x, u in F, and

$$(\left[\begin{smallmatrix} a&&&\\ &a&&\\ &&b&\\ &&&b \end{smallmatrix}\right] f)(u) = \chi(b)\pi(\left[\begin{smallmatrix} a&\\ &b \end{smallmatrix}\right]) f(a^{-1}bu) \tag{128}$$

for u in F and a, b in $F^\times$. The subspace $\mathcal{S}(F^\times, V)$ consisting of functions that vanish at zero is invariant under these actions. We consider the exact sequence

$$0 \longrightarrow \mathcal{S}(F^\times, V) \longrightarrow \mathcal{S}(F, V) \longrightarrow \mathcal{S}(F,V)/\mathcal{S}(F^\times, V) \longrightarrow 0.$$

The quotient $\mathcal{S}(F,V)/\mathcal{S}(F^\times,V)$ is isomorphic to V via the map $f \mapsto f(0)$. The actions of the above subgroups on V are given by

$$\left[\begin{smallmatrix} 1&&&\\ &1&&\\ &x&1&\\ &&&1 \end{smallmatrix}\right] v = \pi(\left[\begin{smallmatrix} 1&\\ x&1 \end{smallmatrix}\right]) v \tag{129}$$

and

$$\left[\begin{smallmatrix} a&&&\\ &a&&\\ &&b&\\ &&&b \end{smallmatrix}\right] v = \chi(b)\pi(\left[\begin{smallmatrix} a&\\ &b \end{smallmatrix}\right]) v. \tag{130}$$

Taking Jacquet modules on the above sequence gives

$$0 \longrightarrow \mathcal{S}(F^\times, V)_{\left[\begin{smallmatrix} 1&&&\\ &1&&\\ &*&1&\\ &&&1 \end{smallmatrix}\right]} \longrightarrow \mathcal{S}(F, V)_{\left[\begin{smallmatrix} 1&&&\\ &1&&\\ &*&1&\\ &&&1 \end{smallmatrix}\right]}$$

$$\longrightarrow \big(\mathcal{S}(F,V)/\mathcal{S}(F^\times,V)\big)_{\left[\begin{smallmatrix}1&&&\\&1&&\\&*&1&\\&&&1\end{smallmatrix}\right]} \longrightarrow 0.$$

In view of (129), the Jacquet module on the right is isomorphic to $V_{\left[\begin{smallmatrix}1&\\ *&1\end{smallmatrix}\right]}$. The action of the diagonal subgroup on $V_{\left[\begin{smallmatrix}1&\\ *&1\end{smallmatrix}\right]}$ is given by the same formula as in (130).

We consider the map from $\mathcal{S}(F^\times,V)$ to itself given by $f\mapsto \big(u\mapsto \pi(\left[\begin{smallmatrix}1&\\&u\end{smallmatrix}\right])f(u)\big)$. This map is an isomorphism of vector spaces. The actions (127) and (128) turn into

$$\Big(\begin{bmatrix}1&&&\\&1&&\\&x&1&\\&&&1\end{bmatrix} f\Big)(u)=\psi(-ux)\pi(\left[\begin{smallmatrix}1&\\ ux&1\end{smallmatrix}\right])f(u) \tag{131}$$

and

$$\Big(\begin{bmatrix}a&&&\\&a&&\\&&b&\\&&&b\end{bmatrix} f\Big)(u)=\chi(b)\pi(\left[\begin{smallmatrix}a&\\&a\end{smallmatrix}\right])f(a^{-1}bu). \tag{132}$$

We define a map p from $\mathcal{S}(F^\times,V)$ to

$$\mathcal{S}(F^\times,V_{\left[\begin{smallmatrix}1&\\ *&1\end{smallmatrix}\right],\psi})=\mathcal{S}(F^\times,V/V(\left[\begin{smallmatrix}1&\\ *&1\end{smallmatrix}\right],\psi))$$

by sending f to f composed with the projection from V to $V/V(\left[\begin{smallmatrix}1&\\ *&1\end{smallmatrix}\right],\psi)$. This map is surjective. The kernel of p consists of all f in $\mathcal{S}(F^\times,V)$ for which there exists a positive integer l such that

$$\int\limits_{\mathfrak{p}^{-l}}\psi(-x)\pi(\left[\begin{smallmatrix}1&\\ x&1\end{smallmatrix}\right])f(u)\,dx=0 \qquad \text{for all } u\in F^\times. \tag{133}$$

Let W be the space of f in $\mathcal{S}(F^\times,V)$ for which there exists a positive integer k such that

$$\int\limits_{\mathfrak{p}^{-k}}\begin{bmatrix}1&&&\\&1&&\\&x&1&\\&&&1\end{bmatrix} f\,dx=0, \tag{134}$$

so that $\mathcal{S}(F^\times,V)/W=\mathcal{S}(F^\times,V)_{\left[\begin{smallmatrix}1&&&\\&1&&\\&*&1&\\&&&1\end{smallmatrix}\right]}$. Let f be in W. The condition (134) means that

$$\int\limits_{\mathfrak{p}^{-k}}\psi(-ux)\pi(\left[\begin{smallmatrix}1&\\ ux&1\end{smallmatrix}\right])f(u)\,dx=0 \qquad \text{for all } u\in F^\times. \tag{135}$$

Since f has compact support in $F^\times$, the conditions (133) and (135) are equivalent. It follows that

$$\mathcal{S}(F^\times,V)_{\left[\begin{smallmatrix}1&&&\\&1&&\\&*&1&\\&&&1\end{smallmatrix}\right]}\cong \mathcal{S}(F^\times,V_{\left[\begin{smallmatrix}1&\\ *&1\end{smallmatrix}\right],\psi}).$$

The diagonal subgroup acts on $\mathcal{S}(F^\times,V_{\left[\begin{smallmatrix}1&\\ *&1\end{smallmatrix}\right],\psi})$ by the same formula as in (132). ∎

6 The main results

Having assembled all the required tools, we are now ready to prove the three main results of this paper mentioned in the introduction.

6.1 Existence of Bessel functionals

In this section we prove that every irreducible, admissible representation (π, V) of $\mathrm{GSp}(4, F)$ which is not a twist of the trivial representation admits a Bessel functional. The proof uses the P_3-module V_{Z^J} and the G^J-module $V_{Z^J,\psi}$. The first module is closely related to the theory of zeta integrals. The second module $V_{Z^J,\psi}$ is the quotient of V by the vector subspace generated by the elements of the form $\pi(\begin{bmatrix} 1&&&z\\ &1&&\\ &&1&\\ &&&1 \end{bmatrix})v - \psi(z)v$ for $v \in V$ and $z \in F$. Evidently, $V_{Z^J,\psi}$ is a G^J module. This module is closely related to the theory of representations of the metaplectic group $\widetilde{\mathrm{SL}}(2, F)$.

6.1.1 Lemma. *Let (π, V) be a smooth representation of N. Then there exists a character θ of N such that $V_{N,\theta} \neq 0$.*

Proof. This follows immediately from Lemma 1.6 of [29]. ∎

Let $\widetilde{\mathrm{SL}}(2, F)$ be the metaplectic group, defined as in Sect. 1 of [29]. Let m be in $F^\times$. We will use the *Weil representation* π_W^m of $\widetilde{\mathrm{SL}}(2, F)$ on $\mathcal{S}(F)$ associated to the quadratic form $q(x) = x^2$ and ψ^m. This is as defined on pp. 3-4 of [39] and p. 223 of [41]. The only explicit property of π_W^m we will use is

$$(\pi_W^m([\begin{smallmatrix} 1&b\\ &1 \end{smallmatrix}], 1)f)(x) = \psi(mbx^2)f(x), \tag{136}$$

for b in F and f in $\mathcal{S}(F)$. We define an action of N_Q, introduced in (6), on the Schwartz space $\mathcal{S}(F)$ by

$$\pi_S^m(\begin{bmatrix} 1&\lambda&\mu&\kappa\\ &1&&\mu\\ &&1&-\lambda\\ &&&1 \end{bmatrix} f)(x) = \psi^m(\kappa + (2x+\lambda)\mu)f(x+\lambda) \tag{137}$$

for f in $\mathcal{S}(F)$. This representation of N_Q is called the *Schrödinger representation.*

Given a smooth, genuine representation (τ, W) of $\widetilde{\mathrm{SL}}(2, F)$, we define a representation τ^J of G^J on the space $W \otimes \mathcal{S}(F)$ by the formulas

$$\tau^J(\begin{bmatrix} 1&&&\\ &a&b&\\ &c&d&\\ &&&1 \end{bmatrix})(v \otimes f) = \tau([\begin{smallmatrix} a&b\\ c&d \end{smallmatrix}], 1)v \otimes \pi_W^m([\begin{smallmatrix} a&b\\ c&d \end{smallmatrix}], 1)f, \tag{138}$$

$$\tau^J(\begin{bmatrix} 1&\lambda&\mu&\kappa\\ &1&&\mu\\ &&1&-\lambda\\ &&&1 \end{bmatrix})(v \otimes f) = v \otimes \pi_S^m(\begin{bmatrix} 1&\lambda&\mu&\kappa\\ &1&&\mu\\ &&1&-\lambda\\ &&&1 \end{bmatrix})f. \tag{139}$$

Computations show that τ^J is a smooth representation of G^J. Moreover, the map that sends τ to τ^J is a bijection between the set of equivalence classes of smooth, genuine representations of $\widetilde{\mathrm{SL}}(2,F)$, and smooth representations of G^J with central character ψ^m. The proof of this fact is based on the Stone-von Neumann Theorem; see Theorem 2.6.2 of [2]. Under this bijection, irreducible τ correspond to irreducible τ^J.

6.1.2 Lemma. *Let m be in $F^\times$. Let (τ^J, W^J) be a non-zero, irreducible, smooth representation of G^J with central character ψ^m. Then $\dim W^J_{N,\theta_{a,0,m}} \leq 1$ for all a in $F^\times$ and $\dim W^J_{N,\theta_{a,0,m}} = 1$ for some a in $F^\times$. This dimension depends only on the class of a in $F^\times/F^{\times 2}$.*

Proof. By the above discussion, there exists an irreducible, genuine, admissible representation τ of $\widetilde{\mathrm{SL}}(2,F)$ such that $\tau^J \cong \tau \otimes \pi^m_{SW}$. Using (136), (137) and iii) of Lemma 5.1.1, an easy calculation shows that

$$W^J_{\left[\begin{smallmatrix}1&&&\\&1&*&*\\&&1&*\\&&&1\end{smallmatrix}\right],\theta_{a,0,m}} \cong W_{\left[\begin{smallmatrix}1&*\\&1\end{smallmatrix}\right],\psi^a}.$$

By Lemme 2 on p. 226 of [41], the space on the right is at most one-dimensional, and is one-dimensional for some a in $F^\times$. Moreover, the dimension depends only on the class of a in $F^\times/F^{\times 2}$. ∎

6.1.3 Proposition. *Let (π, V) be an irreducible, admissible representation of* $\mathrm{GSp}(4,F)$. *Then the following statements are equivalent.*

i) π is not a twist of the trivial representation.

ii) There exists a non-trivial character θ of N such that $V_{N,\theta} \neq 0$.

iii) There exists a non-degenerate character θ of N such that $V_{N,\theta} \neq 0$.

Proof. i) $\Rightarrow$ ii) Assume that $V_{N,\theta} = 0$ for all non-trivial θ. By Lemma 6.1.1, it follows that $V_{N,1} \neq 0$. In particular, the P_3-module V_{Z^J} is non-zero. By using Theorem 3.2.1 and inspecting tables A.5 and A.6 in [28], one can see that V_{Z^J} contains an irreducible subquotient τ of the form $\tau^{P_3}_{\mathrm{GL}(0)}(1)$, or $\tau^{P_3}_{\mathrm{GL}(1)}(\chi)$ for a character χ of $F^\times$, or $\tau^{P_3}_{\mathrm{GL}(2)}(\rho)$ for an irreducible, admissible, infinite-dimensional representation ρ of $\mathrm{GL}(2,F)$; it is here that we use the hypothesis that π is not one-dimensional. For a, b in F we define a character of the subgroup $\left[\begin{smallmatrix}1&*&*\\&1&\\&&1\end{smallmatrix}\right]$ of P_3 by

$$\theta_{a,b}(\left[\begin{smallmatrix}1&x&y\\&1&\\&&1\end{smallmatrix}\right]) = \psi(ax+by). \tag{140}$$

By Lemma 2.5.4 or Lemma 2.5.5 of [28], or the infinite-dimensionality of ρ if $\tau = \tau^{P_3}_{\mathrm{GL}(2)}(\rho)$,

$$\tau_{\left[\begin{smallmatrix}1&*&*\\&1&*\\&&1\end{smallmatrix}\right],\theta_{a,b}} \neq 0$$

for some $(a,b) \neq (0,0)$. This implies that $V_{N,\theta_{a,b,0}} \neq 0$, contradicting our assumption.
ii) $\Rightarrow$ iii) The hypothesis implies that V_{Z^J,ψ^m} is non-zero for some m in $F^\times$. We observe that V_{Z^J,ψ^m} is a smooth G^J representation. By Lemma 2.6 of [3], there exists an irreducible subquotient (τ^J, W^J) of this G^J module. By Lemma 6.1.2, we have $\dim W^J_{N,\theta_{a,0,m}} = 1$ for some a in $F^\times$. This implies that $V_{N,\theta_{a,0,m}} \neq 0$.
iii) $\Rightarrow$ i) is obvious. ∎

6.1.4 THEOREM. *Let (π, V) be an irreducible, admissible representation of* $\mathrm{GSp}(4,F)$. *Assume that π is not one-dimensional. Then π admits a (Λ,θ)-Bessel functional for some non-degenerate character θ of N and some character Λ of T. If π is non-generic and supercuspidal, then every Bessel functional for π is non-split.*

Proof. By Proposition 6.1.3, there exists a non-degenerate θ such that $V_{N,\theta} \neq 0$. Assume that θ is non-split. Then, since the center $F^\times$ of $\mathrm{GSp}(4,F)$ acts by a character on $V_{N,\theta}$ and $T/F^\times$ is compact, $V_{N,\theta}$ decomposes as a direct sum over characters of T. It follows that a (Λ,θ)-Bessel functional exists for some character Λ of T.
Now assume that θ is split. We may assume that S is the matrix in (22). Let V_0, V_1, V_2 be the modules appearing in the P_3-filtration, as in Theorem 3.2.1. Since $V_{N,\theta} \neq 0$, we must have

$$(V_0/V_1)_{\left[\begin{smallmatrix}1&*&*\\&1&*\\&&1\end{smallmatrix}\right],\theta_{0,1}} \neq 0, \quad (V_1/V_2)_{\left[\begin{smallmatrix}1&*&*\\&1&*\\&&1\end{smallmatrix}\right],\theta_{0,1}} \neq 0, \quad \text{or} \quad (V_2)_{\left[\begin{smallmatrix}1&*&*\\&1&*\\&&1\end{smallmatrix}\right],\theta_{0,1}} \neq 0,$$

where we use the notation (140). It is immediate from (38) that the first space is zero. If the second space is non-zero, then π admits a split Bessel functional by iii) of Proposition 3.5.1. If the third space is non-zero, then π is generic by Theorem 3.2.1, and hence, by Proposition 3.4.2, admits a split Bessel functional.
For the last statement, assume that π is non-generic and supercuspidal. Then $V_{Z^J} = 0$ by Theorem 3.2.1. Hence, $V_{N,\theta} = 0$ for any split θ. It follows that all Bessel functionals for π are non-split. ∎

6.2 The table of Bessel functionals

In this section, given a non-supercuspidal representation π, or a π that is in an L-packet with a non-supercuspidal representation, we determine the set of (Λ,θ) for which π admits a (Λ,θ)-Bessel functional.

6.2.1 Lemma. *Let θ be as in (15), and let T be the corresponding torus. Assume that the associated quadratic extension L is a field. Let V_1, V_2, V_3 and W be smooth representations of T. Assume that these four representations all have the same central character. Assume further that there is an exact sequence of T-modules*

$$0 \longrightarrow V_1 \longrightarrow V_2 \longrightarrow V_3 \longrightarrow 0.$$

Then the sequence of T-modules

$$0 \longrightarrow \mathrm{Hom}_T(V_3, W) \longrightarrow \mathrm{Hom}_T(V_2, W) \longrightarrow \mathrm{Hom}_T(V_1, W) \longrightarrow 0$$

is exact.

Proof. It is easy to see that the sequence

$$0 \longrightarrow \mathrm{Hom}_T(V_3, W) \longrightarrow \mathrm{Hom}_T(V_2, W) \longrightarrow \mathrm{Hom}_T(V_1, W)$$

is exact. We will prove the surjectivity of the last map. Let f be in $\mathrm{Hom}_T(V_1, W)$. We extend f to a linear map f_1 from V_2 to W. We define another linear map f_2 from V_2 to W by

$$f_2(v) = \int\limits_{T/F^\times} t^{-1} \cdot f_1(t \cdot v)\, dt.$$

This is well-defined by the condition on the central characters, the compactness of $T/F^\times$, and the smoothness hypothesis. Evidently, f_2 is in $\mathrm{Hom}_T(V_2, W)$ and maps to a multiple of f. ∎

6.2.2 Theorem. *The following table shows the Bessel functionals admitted by the irreducible, admissible, non-supercuspidal representations of* $\mathrm{GSp}(4, F)$. *The column "$L \leftrightarrow \xi$" indicates that the field L is the quadratic extension of F corresponding to the non-trivial, quadratic character ξ of $F^\times$; this is only relevant for representations in groups V and IX. The pairs of characters (χ_1, χ_2) in the "$L = F \times F$" column for types IIIb and IVc refer to the characters of $T = \{\mathrm{diag}(a, b, a, b) : \ a, b \in F^\times\}$ given by $\mathrm{diag}(a, b, a, b) \mapsto \chi_1(a)\chi_2(b)$. In representations of group IX, the symbol μ denotes a non-Galois-invariant character of $L^\times$, where L is the quadratic extension corresponding to ξ. The Galois conjugate of μ is denoted by μ'. The irreducible, admissible, supercuspidal representation of* $\mathrm{GL}(2, F)$ *corresponding to μ is denoted by $\pi(\mu)$. Finally, the symbol N in the table stands for the norm map $\mathrm{N}_{L/F}$. In the split case, the character $\sigma \circ \mathrm{N}$ is the same as (σ, σ). In the table, the phrase "all Λ" means all characters Λ of T whose restriction to $F^\times$ is the central character of the representation of* $\mathrm{GSp}(4, F)$.

		representation	(Λ,θ)-Bessel functional exists exactly for ...		
			$L = F \times F$	L/F a field extension	
				$L \leftrightarrow \xi$	$L \not\leftrightarrow \xi$
I		$\chi_1 \times \chi_2 \rtimes \sigma$ (irred.)	all Λ	all Λ	
II	a	$\chi\mathrm{St}_{\mathrm{GL}(2)} \rtimes \sigma$	all Λ	$\Lambda \neq (\chi\sigma) \circ \mathrm{N}$	
	b	$\chi 1_{\mathrm{GL}(2)} \rtimes \sigma$	$\Lambda = (\chi\sigma) \circ \mathrm{N}$	$\Lambda = (\chi\sigma) \circ \mathrm{N}$	
III	a	$\chi \rtimes \sigma\mathrm{St}_{\mathrm{GSp}(2)}$	all Λ	all Λ	
	b	$\chi \rtimes \sigma 1_{\mathrm{GSp}(2)}$	$\Lambda \in \{(\chi\sigma,\sigma),(\sigma,\chi\sigma)\}$	—	
IV	a	$\sigma\mathrm{St}_{\mathrm{GSp}(4)}$	all Λ	$\Lambda \neq \sigma \circ \mathrm{N}$	
	b	$L(\nu^2, \nu^{-1}\sigma\mathrm{St}_{\mathrm{GSp}(2)})$	$\Lambda = \sigma \circ \mathrm{N}$	$\Lambda = \sigma \circ \mathrm{N}$	
	c	$L(\nu^{3/2}\mathrm{St}_{\mathrm{GL}(2)}, \nu^{-3/2}\sigma)$	$\Lambda = (\nu^{\pm 1}\sigma, \nu^{\mp 1}\sigma)$	—	
	d	$\sigma 1_{\mathrm{GSp}(4)}$	—	—	
V	a	$\delta([\xi,\nu\xi], \nu^{-1/2}\sigma)$	all Λ	$\Lambda \neq \sigma \circ \mathrm{N}$	$\sigma \circ \mathrm{N} \neq \Lambda \neq (\xi\sigma) \circ \mathrm{N}$
	b	$L(\nu^{1/2}\xi\mathrm{St}_{\mathrm{GL}(2)}, \nu^{-1/2}\sigma)$	$\Lambda = \sigma \circ \mathrm{N}$	—	$\Lambda = \sigma \circ \mathrm{N}$
	c	$L(\nu^{1/2}\xi\mathrm{St}_{\mathrm{GL}(2)}, \xi\nu^{-1/2}\sigma)$	$\Lambda = (\xi\sigma) \circ \mathrm{N}$	—	$\Lambda = (\xi\sigma) \circ \mathrm{N}$
	d	$L(\nu\xi, \xi \rtimes \nu^{-1/2}\sigma)$	—	$\Lambda = \sigma \circ \mathrm{N}$	—
VI	a	$\tau(S, \nu^{-1/2}\sigma)$	all Λ	$\Lambda \neq \sigma \circ \mathrm{N}$	
	b	$\tau(T, \nu^{-1/2}\sigma)$	—	$\Lambda = \sigma \circ \mathrm{N}$	
	c	$L(\nu^{1/2}\mathrm{St}_{\mathrm{GL}(2)}, \nu^{-1/2}\sigma)$	$\Lambda = \sigma \circ \mathrm{N}$	—	
	d	$L(\nu, 1_{F^\times} \rtimes \nu^{-1/2}\sigma)$	$\Lambda = \sigma \circ \mathrm{N}$	—	
VII		$\chi \rtimes \pi$	all Λ	all Λ	
VIII	a	$\tau(S, \pi)$	all Λ	$\mathrm{Hom}_T(\pi, \mathbb{C}_\Lambda) \neq 0$	
	b	$\tau(T, \pi)$	—	$\mathrm{Hom}_T(\pi, \mathbb{C}_\Lambda) = 0$	
IX	a	$\delta(\nu\xi, \nu^{-1/2}\pi(\mu))$	all Λ	$\mu \neq \Lambda \neq \mu'$	all Λ
	b	$L(\nu\xi, \nu^{-1/2}\pi(\mu))$	—	$\Lambda = \mu$ or $\Lambda = \mu'$	—
X		$\pi \rtimes \sigma$	all Λ	$\mathrm{Hom}_T(\sigma\pi, \mathbb{C}_\Lambda) \neq 0$	
XI	a	$\delta(\nu^{1/2}\pi, \nu^{-1/2}\sigma)$	all Λ	$\Lambda \neq \sigma \circ \mathrm{N}$ and $\mathrm{Hom}_T(\sigma\pi, \mathbb{C}_\Lambda) \neq 0$	
	b	$L(\nu^{1/2}\pi, \nu^{-1/2}\sigma)$	$\Lambda = \sigma \circ \mathrm{N}$	$\Lambda = \sigma \circ \mathrm{N}$ and $\mathrm{Hom}_T(\pi, \mathbb{C}_1) \neq 0$	
Va*		$\delta^*([\xi,\nu\xi], \nu^{-1/2}\sigma)$	—	$\Lambda = \sigma \circ \mathrm{N}$	—
XIa*		$\delta^*(\nu^{1/2}\pi, \nu^{-1/2}\sigma)$	—	$\Lambda = \sigma \circ \mathrm{N}$ and $\mathrm{Hom}_T(\pi^{\mathrm{JL}}, \mathbb{C}_1) \neq 0$	

Proof. We will go through all representations in the table and explain how the statements follow from our preparatory sections.
I: This follows from Proposition 3.4.2 and Lemma 5.3.2.
IIa: In the split case this follows from Proposition 3.4.2. In the non-split case it follows from Lemma 5.2.2 together with (36).
IIb: This follows from Lemma 5.2.2; see (113).
IIIa: This follows from Proposition 3.4.2 and Lemma 5.3.2.
IIIb: It follows from Lemma 5.3.2 that IIIb type representations have no non-split Bessel functionals. The split case follows from either Proposition 3.5.1 or i) of Lemma 5.3.4. Note that the characters $(\chi\sigma, \sigma)$ and $(\sigma, \chi\sigma)$ are Galois conjugates of each other.
IVd: It is easy to see that the twisted Jacquet modules of the trivial representation are zero.
IVb: By (2.9) of [28] there is a short exact sequence

$$0 \longrightarrow \mathrm{IVb} \longrightarrow \nu^{3/2} 1_{\mathrm{GL}(2)} \rtimes \nu^{-3/2}\sigma \longrightarrow \sigma 1_{\mathrm{GSp}(4)} \longrightarrow 0.$$

Taking twisted Jacquet modules and observing (113), we get

$$(\mathrm{IVb})_{N,\theta} \cong (\nu^{3/2} 1_{\mathrm{GL}(2)} \rtimes \nu^{-3/2}\sigma)_{N,\theta} = \mathbb{C}_{\sigma\circ \mathrm{N}_{L/F}}$$

as T-modules.
IVc: By (2.9) of [28] there is a short exact sequence

$$0 \longrightarrow \mathrm{IVc} \longrightarrow \nu^2 \rtimes \nu^{-1}\sigma 1_{\mathrm{GSp}(2)} \longrightarrow \sigma 1_{\mathrm{GSp}(4)} \longrightarrow 0.$$

Taking twisted Jacquet modules gives

$$(\mathrm{IVc})_{N,\theta} \cong (\nu^2 \rtimes \nu^{-1}\sigma 1_{\mathrm{GSp}(2)})_{N,\theta}.$$

Hence IVc admits the same Bessel functionals as the full induced representation $\nu^2 \rtimes \nu^{-1}\sigma 1_{\mathrm{GSp}(2)}$. By Lemma 5.3.2, any such Bessel functional is necessarily split. Assume that θ is as in (23). Then, using Lemma 5.3.4, it follows that IVc admits the (Λ, θ)-Bessel functional for

$$\Lambda(\begin{bmatrix} a &&& \\ & b && \\ && a & \\ &&& b \end{bmatrix}) = \nu(ab^{-1})\sigma(ab), \tag{141}$$

which we write as $(\nu\sigma, \nu^{-1}\sigma)$. By (29), IVc also admits a (Λ, θ)-Bessel functional for $\Lambda = (\nu^{-1}\sigma, \nu\sigma)$. Again by Lemma 5.3.4, IVc does not admit a (Λ, θ)-Bessel functional for any other Λ.
IVa: In the split case this follows from Proposition 3.4.2. Assume θ is non-split. By (2.9) of [28], there is an exact sequence

$$0 \longrightarrow \sigma \mathrm{St}_{\mathrm{GSp}(4)} \longrightarrow \nu^2 \rtimes \nu^{-1}\sigma \mathrm{St}_{\mathrm{GSp}(2)} \longrightarrow \mathrm{IVb} \longrightarrow 0.$$

Taking Jacquet modules, we get

$$0 \longrightarrow (\sigma \mathrm{St}_{\mathrm{GSp}(4)})_{N,\theta} \longrightarrow (\nu^2 \rtimes \nu^{-1}\sigma \mathrm{St}_{\mathrm{GSp}(2)})_{N,\theta} \longrightarrow (\mathrm{IVb})_{N,\theta} \longrightarrow 0.$$

Keeping in mind Lemma 6.2.1, the result now follows from Lemma 5.3.2 and the result for IVb.
Vd: This was proved in Corollary 4.7.2.
Vb and Vc: Let ξ be a non-trivial quadratic character of $F^\times$. By (2.10) of [28], there are exact sequences

$$0 \longrightarrow \mathrm{Vb} \longrightarrow \nu^{1/2}\xi 1_{\mathrm{GL}(2)} \rtimes \xi\nu^{-1/2}\sigma \longrightarrow \mathrm{Vd} \longrightarrow 0$$

and

$$0 \longrightarrow \mathrm{Vc} \longrightarrow \nu^{1/2}\xi 1_{\mathrm{GL}(2)} \rtimes \nu^{-1/2}\sigma \longrightarrow \mathrm{Vd} \longrightarrow 0.$$

Taking Jacquet modules and observing (113), we get

$$0 \longrightarrow (\mathrm{Vb})_{N,\theta} \longrightarrow \mathbb{C}_{\sigma\circ\mathrm{N}_{L/F}} \longrightarrow (\mathrm{Vd})_{N,\theta} \longrightarrow 0 \tag{142}$$

and

$$0 \longrightarrow (\mathrm{Vc})_{N,\theta} \longrightarrow \mathbb{C}_{(\xi\sigma)\circ\mathrm{N}_{L/F}} \longrightarrow (\mathrm{Vd})_{N,\theta} \longrightarrow 0. \tag{143}$$

Hence the results for Vb and Vc follow from the result for Vd.
Va: In the split case this follows from Proposition 3.4.2. Assume θ is non-split. Assume first that ξ corresponds to the quadratic extension L/F. As we just saw, $(\mathrm{Vb})_{N,\theta} = 0$ in this case. By (2.10) of [28], there is an exact sequence

$$0 \longrightarrow \mathrm{Va} \longrightarrow \nu^{1/2}\xi\mathrm{St}_{\mathrm{GL}(2)} \rtimes \nu^{-1/2}\sigma \longrightarrow \mathrm{Vb} \longrightarrow 0.$$

Taking Jacquet modules, it follows that

$$(\mathrm{Va})_{N,\theta} = (\nu^{1/2}\xi\mathrm{St}_{\mathrm{GL}(2)} \rtimes \nu^{-1/2}\sigma)_{N,\theta}.$$

By Lemma 5.2.2, the space of (Λ,θ)-Bessel functionals on the representation Va is isomorphic to $\mathrm{Hom}_T(\sigma\xi\mathrm{St}_{\mathrm{GL}(2)}, \mathbb{C}_\Lambda)$. Using (36), it follows that Va admits a (Λ,θ)-Bessel functional if and only if $\Lambda \neq (\sigma\xi)\circ\mathrm{N}_{L/F} = \sigma\circ\mathrm{N}_{L/F}$.
Now assume that ξ does not correspond to the quadratic extension L/F. Then, by what we already proved for Vb, we have an exact sequence

$$0 \longrightarrow (\mathrm{Va})_{N,\theta} \longrightarrow (\nu^{1/2}\xi\mathrm{St}_{\mathrm{GL}(2)} \rtimes \nu^{-1/2}\sigma)_{N,\theta} \longrightarrow \mathbb{C}_{\sigma\circ\mathrm{N}_{L/F}} \longrightarrow 0. \tag{144}$$

Using Lemma 6.2.1, it follows that the possible characters Λ for Va are those of $(\nu^{1/2}\xi\mathrm{St}_{\mathrm{GL}(2)} \rtimes \nu^{-1/2}\sigma)_{N,\theta}$ with the exception of $\sigma\circ\mathrm{N}_{L/F}$. By Lemma 5.2.2 and (36), these are all characters other than $\sigma\circ\mathrm{N}_{L/F}$ and $(\xi\sigma)\circ\mathrm{N}_{L/F}$.
VIc and VId: By (2.11) of [28], there is an exact sequence

$$0 \longrightarrow \mathrm{VIc} \longrightarrow 1_{F^\times} \rtimes \sigma 1_{\mathrm{GSp}(2)} \longrightarrow \mathrm{VId} \longrightarrow 0.$$

It follows from Lemma 5.3.2 that VIc and VId have no non-split Bessel functionals. The split case follows from Proposition 3.5.1.
VIa: In the split case this follows from Proposition 3.4.2. Assume that θ is non-split. By (2.11) of [28], there is an exact sequence

$$0 \longrightarrow \mathrm{VIa} \longrightarrow \nu^{1/2}\mathrm{St}_{\mathrm{GL}(2)} \rtimes \nu^{-1/2}\sigma \longrightarrow \mathrm{VIc} \longrightarrow 0. \tag{145}$$

Taking Jacquet modules and observing the result for VIc, we get $(\mathrm{VIa})_{N,\theta} = (\nu^{1/2}\mathrm{St}_{\mathrm{GL}(2)} \rtimes \nu^{-1/2}\sigma)_{N,\theta}$. Hence the result follows from Lemma 5.2.2 and (36).

VIb: By (2.11) of [28], there is an exact sequence

$$0 \longrightarrow (\mathrm{VIb})_{N,\theta} \longrightarrow (\nu^{1/2}1_{\mathrm{GL}(2)} \rtimes \nu^{-1/2}\sigma)_{N,\theta} \longrightarrow (\mathrm{VId})_{N,\theta} \longrightarrow 0. \tag{146}$$

By (113), the middle term equals $\mathbb{C}_{\sigma\circ \mathrm{N}_{L/F}}$. One-dimensionality implies that the sequence splits, so that

$$\mathrm{Hom}_T(\mathbb{C}_{\sigma\circ \mathrm{N}_{L/F}}, \mathbb{C}_\Lambda) = \mathrm{Hom}_D(\mathrm{VIb}, \mathbb{C}_{\Lambda\otimes\theta}) \oplus \mathrm{Hom}_D(\mathrm{VId}, \mathbb{C}_{\Lambda\otimes\theta}) \tag{147}$$

(D is the Bessel subgroup defined in (17)). Hence the VIb case follows from the known result for VId.

VII: This follows from Proposition 3.4.2 and Lemma 5.3.2.

VIIIa and VIIIb: In the split case this follows from Proposition 3.4.2 and v) of Proposition 3.5.1. Assume that θ is non-split. Since we are in a unitarizable situation, the sequence

$$0 \longrightarrow \mathrm{VIIIa} \longrightarrow 1_{F^\times} \rtimes \pi \longrightarrow \mathrm{VIIIb} \longrightarrow 0$$

splits. It follows that

$$\mathrm{Hom}_D(1_{F^\times} \rtimes \pi, \mathbb{C}_{\Lambda\otimes\theta}) = \mathrm{Hom}_D(\mathrm{VIIIa}, \mathbb{C}_{\Lambda\otimes\theta}) \oplus \mathrm{Hom}_D(\mathrm{VIIIb}, \mathbb{C}_{\Lambda\otimes\theta}). \tag{148}$$

By Lemma 5.3.2, the space on the left is one-dimensional for any Λ. Therefore the Bessel functionals of VIIIb are complementary to those of VIIIa.

Assume that VIIIa admits a (Λ,θ)-Bessel functional. Then, by Corollary 4.7.1 and Theorem 4.6.3, we have $\mathrm{Hom}_T(\pi, \mathbb{C}_\Lambda) \neq 0$. Conversely, assume that $\mathrm{Hom}_T(\pi, \mathbb{C}_\Lambda) \neq 0$ and assume that VIIIa does not admit a (Λ,θ)-Bessel functional; we will obtain a contradiction. By (148), we have $\mathrm{Hom}_D(\mathrm{VIIIb}, \mathbb{C}_{\Lambda\otimes\theta}) \neq 0$. By Corollary 4.7.1 and Theorem 4.6.3, we have $\mathrm{Hom}_T(\pi^{\mathrm{JL}}, \mathbb{C}_\Lambda) \neq 0$. This contradicts (34).

The result for VIIIb now follows from (148).

IXb: This was proved in Corollary 4.7.2.

IXa: In the split case this follows from Proposition 3.4.2. Assume that θ is non-split. We have an exact sequence

$$0 \longrightarrow \mathrm{IXa} \longrightarrow \nu\xi \rtimes \nu^{-1/2}\pi \longrightarrow \mathrm{IXb} \longrightarrow 0.$$

By Lemma 5.3.2, the space $\mathrm{Hom}_D(\nu\xi \rtimes \nu^{-1/2}\pi, \mathbb{C}_{\Lambda\otimes\theta})$ is one-dimensional, for any character Λ of $L^\times$ satisfying the central character condition. It follows that the possible Bessel functionals of IXa are complementary to those of IXb.

X: In the split case this follows from Proposition 3.4.2. In the non-split case it follows from Lemma 5.2.2.

XIa and XIb: In the split case this follows from Proposition 3.4.2 and Proposition 3.5.1; note that the V_1/V_2 quotient of XIb equals $\tau^{P_3}_{\mathrm{GL}(1)}(\nu\sigma)$ by Table A.6 of [28]. Assume that L/F is not split, and consider the exact sequence

$$0 \longrightarrow (\mathrm{XIa})_{N,\theta} \longrightarrow (\nu^{1/2}\pi \rtimes \nu^{-1/2}\sigma)_{N,\theta} \longrightarrow (\mathrm{XIb})_{N,\theta} \longrightarrow 0. \tag{149}$$

It follows from Lemma 6.2.1 that

$$\mathrm{Hom}_D(\nu^{1/2}\pi \rtimes \nu^{-1/2}\sigma, \mathbb{C}_{\Lambda\otimes\theta}) = \mathrm{Hom}_D(\mathrm{XIa}, \mathbb{C}_{\Lambda\otimes\theta}) \oplus \mathrm{Hom}_D(\mathrm{XIb}, \mathbb{C}_{\Lambda\otimes\theta}). \tag{150}$$

Observe here that, by Lemma 5.2.2, the left side equals $\mathrm{Hom}_T(\sigma\pi, \mathbb{C}_\Lambda)$, which is at most one-dimensional.

Assume that the representation XIa admits a (Λ, θ)-Bessel functional. Then $\Lambda \neq \sigma \circ \mathrm{N}_{L/F}$ and $\mathrm{Hom}_T(\sigma\pi, \mathbb{C}_\Lambda) \neq 0$ by Corollary 4.7.1 and Theorem 4.6.3. Conversely, assume that $\Lambda \neq \sigma \circ \mathrm{N}_{L/F}$ and $\mathrm{Hom}_T(\sigma\pi, \mathbb{C}_\Lambda) \neq 0$. Assume also that XIa does not admit a (Λ, θ)-Bessel functional; we will obtain a contradiction. By the one-dimensionality of the space on the left hand side of (150), we have $\mathrm{Hom}_D(\mathrm{XIb}, \mathbb{C}_{\Lambda\otimes\theta}) \neq 0$. By Corollary 4.7.1 and Theorem 4.6.3, we conclude $\Lambda = \sigma \circ \mathrm{N}_{L/F}$, contradicting our assumption.

Assume that the representation XIb admits a (Λ, θ)-Bessel functional. Then $\Lambda = \sigma \circ \mathrm{N}_{L/F}$ and $\mathrm{Hom}_T(\pi, \mathbb{C}_1) \neq 0$ by Corollary 4.7.1 and Theorem 4.6.3. Conversely, assume that $\Lambda = \sigma \circ \mathrm{N}_{L/F}$ and $\mathrm{Hom}_T(\pi, \mathbb{C}_1) \neq 0$. Assume also that XIb does not admit a (Λ, θ)-Bessel functional; we will obtain a contradiction. By our assumption, the space on the left hand side of (150) is one-dimensional. Hence $\mathrm{Hom}_D(\mathrm{XIa}, \mathbb{C}_{\Lambda\otimes\theta}) \neq 0$. By what we have already proven, this implies $\Lambda \neq \sigma \circ \mathrm{N}_{L/F}$, a contradiction.

<u>Va*</u>: This was proved in Corollary 4.7.2.

<u>XIa*</u>: By Proposition 3.5.1, the representation XIa* has no split Bessel functionals. Assume that θ is non-split. By Corollary 4.7.1 and Theorem 4.6.3, if XIa* admits a (Λ, θ)-Bessel functional, then $\Lambda = \sigma \circ \mathrm{N}$ and $\mathrm{Hom}_T(\pi^{\mathrm{JL}}, \mathbb{C}_1) \neq 0$. Conversely, assume that $\Lambda = \sigma \circ \mathrm{N}$ and $\mathrm{Hom}_T(\pi^{\mathrm{JL}}, \mathbb{C}_1) \neq 0$. By Corollary 4.7.3, the twisted Jacquet module $\delta^*(\nu^{1/2}\pi, \nu^{-1/2}\sigma)_{N,\theta}$ is one-dimensional. Therefore, XIa* does admit a (Λ', θ)-Bessel functional for some Λ'. By what we already proved, $\Lambda' = \Lambda$.

This concludes the proof. ∎

6.3 Some cases of uniqueness

Let (π, V) be an irreducible, admissible representation of $\mathrm{GSp}(4, F)$. Using the notations from Sect. 2.4, consider (Λ, θ)-Bessel functionals for π. We say that such functionals are *unique* if the dimension of the space $\mathrm{Hom}_D(V, \mathbb{C}_{\Lambda\otimes\theta})$ is at most 1. In this section we will prove the uniqueness of split Bessel functionals for all representations, and the uniqueness of non-split Bessel functionals for all non-supercuspidal representations.

As far as we know, a complete proof that Bessel functionals are unique for all (Λ, θ) and all representations π has not yet appeared in the literature. In [17] it is proved that $(1, \theta)$-Bessel functionals are unique if π has trivial central character. The main ingredient for this proof is Theorem 1' of [10]. In [15] it is proved that (Λ, θ)-Bessel functionals are unique if π has trivial central character. The proof is based on a generalization of Theorem 1' of [10]. In [31] it is stated, without proof, that (Λ, θ)-Bessel functionals are unique if π

is supercuspidal and has trivial central character. In [23] it is remarked that the uniqueness of (Λ,θ)-Bessel functionals in the general case can be proven by extending the arguments of [17] and [15], though a proof of this is not given in [23].

6.3.1 Lemma. *Let σ_1 be a character of $F^\times$, and let (π_1,V_1) be an irreducible, admissible representation of* $\mathrm{GL}(2,F)$. *Let the matrix S be as in* (22), *and θ be as in* (24). *The resulting group T is then given by* (24). *Let (π,V) be an irreducible, admissible representation of* $\mathrm{GSp}(4,F)$. *Assume there is an exact sequence*

$$\pi_1\rtimes\sigma_1\longrightarrow\pi\longrightarrow 0. \tag{151}$$

Let Λ be a character of T. If Λ is not equal to one of the characters Λ_1 or Λ_2, given by

$$\Lambda_1(\mathrm{diag}(a,b,a,b))=\nu^{1/2}(a)\nu^{-1/2}(b)\sigma_1(ab)\omega_{\pi_1}(a), \tag{152}$$

$$\Lambda_2(\mathrm{diag}(a,b,a,b))=\nu^{-1/2}(a)\nu^{1/2}(b)\sigma_1(ab)\omega_{\pi_1}(b), \tag{153}$$

then (Λ,θ)-Bessel functionals are unique.

Proof. Since π is a quotient of $\pi_1\rtimes\sigma_1$, it suffices to prove that $\mathrm{Hom}_D(\pi_1\rtimes\sigma_1,\mathbb{C}_{\Lambda\otimes\theta})$ is at most one-dimensional. Any element β of this space factors through the Jacquet module $(\pi_1\rtimes\sigma_1)_{N,\theta}$. These Jacquet modules were calculated in Lemma 5.2.2 ii). Using the notation of this lemma, the assumption about Λ implies that restriction of β to J_2 establishes an injection

$$\mathrm{Hom}_D(\pi_1\rtimes\sigma_1,\mathbb{C}_{\Lambda\otimes\theta})\longrightarrow\mathrm{Hom}_{\left[\begin{smallmatrix}*&\\&*\end{smallmatrix}\right]}(\sigma_1\pi_1,\mathbb{C}_\Lambda).$$

The space on the right is at most one-dimensional; see Sect. 5.2. This proves our statement. ■

6.3.2 Theorem. *Let (π,V) be an irreducible, admissible representation of* $\mathrm{GSp}(4,F)$.

i) *Split Bessel functionals for π are unique.*

ii) *Non-split Bessel functionals for π are unique, if π is not supercuspidal, or if π is of type* Va* *or* XIa*.

Proof. i) By Proposition 3.5.1, we may assume that π is generic. Let the matrix S be as in (22), and θ be as in (23). The resulting group T is then given by (24). Let Λ be a character of T. We use the fact that any (Λ,θ)-Bessel functional β on V factors through the P_3-module V_{Z^J}.
Assume that π is supercuspidal. Then, by Theorem 3.2.1, the associated P_3-module V_{Z^J} equals $\tau^{P_3}_{\mathrm{GL}(0)}(1)$. Therefore, the space of (Λ,θ)-Bessel functionals

on V equals the space of linear functionals considered in Lemma 2.5.4 of [28]. By this lemma, this space is one-dimensional.

Now assume that π is non-supercuspidal. As in the proof of Proposition 3.5.1, we write the semisimplification of the quotient V_1/V_2 in the P_3-filtration as $\sum_{i=1}^n \tau_{\mathrm{GL}(1)}^{P_3}(\chi_i)$ with characters χ_i of $F^\times$. Let $C(\pi)$ be the set of characters χ_i. Proposition 2.5.7 of [28] states that if the character $a \mapsto \Lambda(\mathrm{diag}(a,1,a,1))$ is not contained in the set $\nu^{-1}C(\pi)$, then the set of (Λ,θ)-Bessel functionals is at most one-dimensional (note that the arguments in the proof of this proposition do not require the hypothesis of trivial central character). The table below lists the sets $\nu^{-1}C(\pi)$ for all generic non-supercuspidal representations. This table implies that (Λ,θ)-Bessel functionals for types VII, VIIIa and IXa are unique. Assume that π is not one these types. Then there exists a sequence as in (151) for some irreducible, admissible representation π_1 of $\mathrm{GL}(2,F)$ and some character σ_1 of $F^\times$. These π_1 and σ_1 are listed in the table below. Let Λ_1, Λ_2 be the characters defined in (152) and (153). Note that, since Λ_1 and Λ_2 are Galois conjugate, we have

$$\dim \mathrm{Hom}_D(\pi, \mathbb{C}_{\Lambda_1\otimes\theta}) = \dim \mathrm{Hom}_D(\pi, \mathbb{C}_{\Lambda_2\otimes\theta}) \tag{154}$$

by (29). By Lemma 6.3.1, it suffices to prove that these spaces are one-dimensional. Define characters λ_1, λ_2 of $F^\times$ by

$$\lambda_1(a) = \Lambda_1(\mathrm{diag}(a,1,a,1)) = \nu^{1/2}(a)\sigma_1(a)\omega_{\pi_1}(a),$$
$$\lambda_2(a) = \Lambda_2(\mathrm{diag}(a,1,a,1)) = \nu^{-1/2}(a)\sigma_1(a).$$

The set $\{\lambda_1,\lambda_2\}$ is listed in the table below for each representation. By the previous paragraph, the spaces (154) are one-dimensional if $\{\lambda_1,\lambda_2\}$ is not a subset of $\nu^{-1}C(\pi)$. This can easily be verified using the table below.

π	π_1	σ_1	$\{\lambda_1,\lambda_2\}$	$\nu^{-1}C(\pi)$
I	$\chi_1\times\chi_2$	σ	$\{\nu^{1/2}\chi_1\chi_2\sigma, \nu^{-1/2}\sigma\}$	$\{\nu^{1/2}\chi_1\chi_2\sigma, \nu^{1/2}\chi_1\sigma, \nu^{1/2}\chi_2\sigma, \nu^{1/2}\sigma\}$
IIa	$\chi\mathrm{St}_{\mathrm{GL}(2)}$	σ	$\{\nu^{1/2}\chi^2\sigma, \nu^{-1/2}\sigma\}$	$\{\nu^{1/2}\chi^2\sigma, \nu^{1/2}\sigma, \nu\chi\sigma\}$
IIIa	$\chi^{-1}\times\nu^{-1}$	$\nu^{1/2}\chi\sigma$	$\{\sigma,\chi\sigma\}$	$\{\nu\chi\sigma,\nu\sigma\}$
IVa	$\nu^{-3/2}\mathrm{St}_{\mathrm{GL}(2)}$	$\nu^{3/2}\sigma$	$\{\nu^{-1}\sigma,\nu\sigma\}$	$\{\nu^2\sigma\}$
Va	$\nu^{-1/2}\xi\mathrm{St}_{\mathrm{GL}(2)}$	$\nu^{1/2}\xi\sigma$	$\{\xi\sigma\}$	$\{\nu\sigma,\nu\xi\sigma\}$
VIa	$\nu^{-1/2}\mathrm{St}_{\mathrm{GL}(2)}$	$\nu^{1/2}\sigma$	$\{\sigma\}$	$\{\nu\sigma\}$
VII	—	—	—	$\varnothing$
VIIIa	—	—	—	$\varnothing$
IXa	—	—	—	$\varnothing$
X	π	σ	$\{\nu^{1/2}\omega_\pi\sigma, \nu^{-1/2}\sigma\}$	$\{\nu^{1/2}\omega_\pi\sigma, \nu^{1/2}\sigma\}$
XIa	$\nu^{-1/2}\pi$	$\nu^{1/2}\sigma$	$\{\sigma\}$	$\{\nu\sigma\}$

ii) Assume first that π is not supercuspidal. Then there exist an irreducible, admissible representation π_1 of $\mathrm{GL}(2,F)$ and a character σ of $F^\times$ such that π is either a quotient of $\pi_1 \rtimes \sigma$, or a quotient of $\sigma \rtimes \pi_1$. The assertion of ii) now follows from i) of Lemma 5.2.2 and Lemma 5.3.2.
Now assume that $\pi = \delta^*([\xi, \nu\xi], \nu^{-1/2}\sigma)$ is of type Va*. Suppose that $\mathrm{Hom}_D(\pi, \mathbb{C}_{\Lambda\otimes\theta})$ is non-zero for some θ and Λ, with L being a field. By our main result Theorem 6.2.2, the quadratic extension L is the field corresponding to ξ and $\Lambda = \sigma \circ \mathrm{N}_{L/F}$. By Corollary 4.7.3, the Jacquet module $\pi_{N,\theta}$ is one-dimensional. This implies that $\mathrm{Hom}_D(\pi, \mathbb{C}_{\Lambda\otimes\theta})$ is one-dimensional.
Finally, assume that $\pi = \delta^*(\nu^{1/2}\pi, \nu^{-1/2}\sigma)$ is of type XIa*. Suppose that $\mathrm{Hom}_D(\pi, \mathbb{C}_{\Lambda\otimes\theta})$ is non-zero for some θ and Λ, with L being a field. By our main result Theorem 6.2.2, we have $\Lambda = \sigma \circ \mathrm{N}_{L/F}$ and $\mathrm{Hom}_T(\pi^{\mathrm{JL}}, \mathbb{C}_1) \neq 0$. By Corollary 4.7.3, the Jacquet module $\pi_{N,\theta}$ is one-dimensional. This implies that $\mathrm{Hom}_D(\pi, \mathbb{C}_{\Lambda\otimes\theta})$ is one-dimensional. ∎

7 Some applications

We present two applications that result from the methods used in this paper. The first application is a characterization of irreducible, admissible, non-generic representations of $\mathrm{GSp}(4,F)$ in terms of their twisted Jacquet modules and their Fourier-Jacobi quotient. The second application concerns the existence of certain vectors with good invariance properties.

7.1 Characterizations of non-generic representations

As before, we fix a non-trivial character ψ of F.

7.1.1 Lemma. *Let (π, V) be a non-generic, supercuspidal, irreducible, admissible representation of $\mathrm{GSp}(4,F)$. Then $\dim V_{N,\theta} < \infty$ for all non-degenerate θ.*

Proof. If θ is split, then $V_{N,\theta} = 0$ by Theorem 3.2.1. Assume that θ is not split. Let $\theta = \theta_S$ with S as in (11). We may assume that $\dim V_{N,\theta} \neq 0$. Let X be as in (57). By Theorem 5.6 of [8], there exists an irreducible, admissible representation σ of $\mathrm{GO}(X)$ such that $\mathrm{Hom}_R(\omega, \pi \otimes \sigma) \neq 0$; here, ω is the Weil representation defined in Sect. 4.4. By i) of Theorem 4.4.6, the set Ω_S is non-empty. By Proposition 4.4.7, the dimension of $V_{N,\theta}$ is finite. ∎

Let W be a smooth representation of N. We will consider the dimensions of the complex vector spaces $W_{N,\theta_{a,b,c}}$. Fix representatives $a_1, \ldots, a_t$ for $F^\times/F^{\times 2}$. We define

$$d(W) = \sum_{i=1}^{t} \dim W_{N,\theta_{a_i,0,1}}.$$

If $0 = W_0 \subset W_1 \subset W_2 \subset \cdots \subset W_k = W$ is a chain of N subspaces, then

$$d(W) = \sum_{j=1}^{k} d(W_j/W_{j-1}). \tag{155}$$

If one of the spaces $W_{N,\theta_{a_i,0,1}}$ is infinite-dimensional, then this equality still holds in the sense that both sides are infinite.

7.1.2 LEMMA. *Let W^J be a non-zero, irreducible, smooth representation of G^J admitting ψ as a central character. Then $1 \leq d(W^J) \leq \#F^\times/F^{\times 2}$.*

Proof. This follows immediately from Lemma 6.1.2. ■

7.1.3 LEMMA. *Let (τ^J, W^J) be a smooth representation of G^J. Then W^J has finite length if and only if $d(W^J)$ is finite. If it has finite length, then*

$$\mathrm{length}(W^J) \leq d(W^J) \leq \mathrm{length}(W^J) \cdot \#F^\times/F^{\times 2}.$$

Proof. Assume that W^J has finite length. Let

$$0 = W_0 \subset W_1 \subset W_2 \subset \cdots \subset W_k = W^J$$

be a chain of G^J spaces such that W_j/W_{j-1} is not zero and irreducible. By (155), we have

$$d(W^J) = \sum_{j=1}^{k} d(W_j/W_{j-1}).$$

By Lemma 7.1.2, $1 \leq d(W_j/W_{j-1}) \leq \#F^\times/F^{\times 2}$ for $j = 1, \ldots, k$. It follows that $d(W^J)$ is finite, and that the asserted inequalities hold.
If W^J has infinite length, a similar argument shows that $d(W)$ is infinite. ■

7.1.4 THEOREM. *Let (π, V) be an irreducible, admissible representation of $\mathrm{GSp}(4, F)$. The following statements are equivalent.*

i) π *is not generic.*

ii) $\dim V_{N,\theta} < \infty$ *for all split* θ.

iii) $\dim V_{N,\theta} < \infty$ *for all non-degenerate* θ.

iv) *The* G^J*-representation* $V_{Z^J,\psi}$ *has finite length.*

Proof. i) $\Rightarrow$ iii) Assume that π is not generic. Let θ be a non-degenerate character of N. Assume first that θ is split. Then $V_{N,\theta}$ can be calculated from the P_3-filtration of π. As in the proof of Proposition 3.5.1 we see that $V_{N,\theta}$ is finite-dimensional.
Now assume that θ is not split. If π is supercuspidal, then $\dim V_{N,\theta} < \infty$ by Lemma 7.1.1. Assume that π is not supercuspidal. Then the table of Bessel functionals shows that π admits (Λ, θ)-Bessel functionals only for finitely many Λ. Since every Λ can occur in $V_{N,\theta}$ at most once by the uniqueness of Bessel functionals (Theorem 6.3.2), this implies that $V_{N,\theta}$ is finite-dimensional.
iii) $\Rightarrow$ ii) is trivial.
ii) $\Rightarrow$ i) Assume that π is generic. Then the subspace V_2 of the P_3-module V_{Z^J} from Theorem 3.2.1 is non-zero. In fact, this subspace is isomorphic to the representation $\tau_{\mathrm{GL}(0)}^{P_3}(1)$ defined in (37). By Lemma 2.5.4 of [28], the space

$$(V_2)_{\left[\begin{smallmatrix}1&*&*\\&1&\\&&1\end{smallmatrix}\right],\theta_{0,1}},$$

where $\theta_{a,b}$ is defined in (140), is infinite-dimensional. This implies that $V_{N,\theta_{0,1,0}}$ is infinite-dimensional, contradicting the hypothesis in ii).
iii) $\Leftrightarrow$ iv) Let $W^J = V_{Z^J,\psi}$. Then $W^J_{N,\theta_{a,0,1}} = V_{N,\theta_{a,0,1}}$ for any a in $F^\times$, so that $d(W^J) = d(V)$. Lemma 7.1.3 therefore implies that iii) and iv) are equivalent. ∎

For more thoughts on $V_{Z^J,\psi}$, see [1]. Theorem 7.1.4 answers one of the questions mentioned at the end of this paper.

7.2 Invariant vectors

Let (π, V) be an irreducible, admissible representation of $\mathrm{GSp}(4, F)$. In this section we will prove the existence of a vector v in V such that $\mathrm{diag}(1,1,c,c)v = v$ for all units c in the ring of integers $\mathfrak{o}$ of F. This result was motivated by a question of Abhishek Saha; see [33].
Our main tool will be the G^J-module $V_{Z^J,\psi}$ for a smooth representation (π, V) of $\mathrm{GSp}(4, F)$. Throughout this section we will make a convenient assumption about the character ψ of F, namely that ψ has conductor $\mathfrak{o}$. By definition, this means that ψ is trivial on $\mathfrak{o}$, but not on $\mathfrak{p}^{-1}$, where $\mathfrak{p}$ is the maximal ideal of $\mathfrak{o}$. We normalize the Haar measure on F such that $\mathfrak{o}$ has volume 1. Let q be the cardinality of the residue class field $\mathfrak{o}/\mathfrak{p}$.
In this section, we will abbreviate

$$d(c) = \begin{bmatrix}1&&&\\&1&&\\&&c&\\&&&c\end{bmatrix}, \qquad z(x) = \begin{bmatrix}1&&&x\\&1&&\\&&1&\\&&&1\end{bmatrix}$$

for c in $F^\times$ and x in F.

7.2.1 Lemma. *Let (π, V) be a smooth representation of $\mathrm{GSp}(4, F)$. Let $p : V \to V_{Z^J,\psi}$ be the projection map, and let w in $V_{Z^J,\psi}$ be non-zero. Then there* *exists a positive integer m and a non-zero vector v in V with the following properties.*

i) $p(v) = w$.

ii) $\pi(z(x))v = \psi(x)v$ *for all* $x \in \mathfrak{p}^{-m}$.

iii) $\pi(d(c))v = v$ *for all* $c \in 1 + \mathfrak{p}^m$.

Proof. Let v_0 in V be such that $p(v_0) = w$. Let m be a positive integer such that $\pi(d(c))v_0 = v_0$ for all $c \in 1 + \mathfrak{p}^m$. Set $v = q^{-m} \int\limits_{\mathfrak{p}^{-m}} \psi(-x)\pi(z(x))v_0 \, dx$. Then $p(v) = w$. In particular, v is not zero. Evidently, v has property ii). Moreover, for c in $1 + \mathfrak{p}^m$,

$$\begin{aligned}
\pi(d(c))(v) &= q^{-m} \int\limits_{\mathfrak{p}^{-m}} \psi(-x)\pi(z(xc^{-1})d(c))v_0 \, dx \\
&= q^{-m} \int\limits_{\mathfrak{p}^{-m}} \psi(-xc)\pi(z(x))v_0 \, dx \\
&= q^{-m} \int\limits_{\mathfrak{p}^{-m}} \psi(-x)\pi(z(x))v_0 \, dx \\
&= v.
\end{aligned}$$

This concludes the proof. ∎

7.2.2 Lemma. *Let (π, V) be a smooth representation of $\mathrm{GSp}(4, F)$. Let $p : V \to V_{Z^J,\psi}$ be the projection map. Let m be a positive integer. Assume that v in V is such that $\pi(z(x))v = \psi(x)v$ for all $x \in \mathfrak{p}^{-m}$. If c is in $\mathfrak{o}^\times$ but not in $1 + \mathfrak{p}^m$, then $p(\pi(d(c))v) = 0$.*

Proof. Let $w = \pi(d(c))v$. To show that $p(w) = 0$ it is enough to show that

$$\int\limits_{\mathfrak{p}^{-m}} \psi(-x)\pi(z(x))w \, dx = 0$$

because $p(\int_{\mathfrak{p}^{-m}} \psi(-x)\pi(z(x))w \, dx) = \int_{\mathfrak{p}^{-m}} \psi(-x)\psi(x)p(w) \, dx = q^m p(w)$. Indeed,

$$\begin{aligned}
\int\limits_{\mathfrak{p}^{-m}} \psi(-x)\pi(z(x))w \, dx &= \int\limits_{\mathfrak{p}^{-m}} \psi(-x)\pi(z(x)d(c))v \, dx \\
&= \pi(d(c)) \int\limits_{\mathfrak{p}^{-m}} \psi(-x)\pi(z(xc))v \, dx
\end{aligned}$$

$$= \pi(d(c)) \int_{\mathfrak{p}^{-m}} \psi(-x)\psi(xc)v\,dx$$

$$= \Big(\int_{\mathfrak{p}^{-m}} \psi(x(c-1))\,dx\Big)\pi(d(c))v$$

$$= 0,$$

since $c \notin 1+\mathfrak{p}^m$ and ψ has conductor $\mathfrak{o}$. ∎

7.2.3 Proposition. *Let (π, V) be a smooth representation of $\mathrm{GSp}(4,F)$. Let $p : V \to V_{Z^J,\psi}$ be the projection map. Let w be in $V_{Z^J,\psi}$. Then there exists a unique vector v in V with the following properties.*

i) $p(v) = w$.

ii) $\pi(z(x))v = v$ *for all* $x \in \mathfrak{o}$.

iii) $\int\limits_{\mathfrak{p}^{-1}} \pi(z(x))v\,dx = 0$.

iv) $\pi(d(c))v = v$ *for all* $c \in \mathfrak{o}^\times$.

Proof. For the existence part we may assume that w is non-zero. Let the positive integer m and v in V be as in Lemma 7.2.1. Define $v_1 = q^m \int\limits_{\mathfrak{o}^\times} \pi(d(c))v\,dc$. Then, by Lemma 7.2.2,

$$\begin{aligned} p(v_1) &= q^m \int_{\mathfrak{o}^\times} p(\pi(d(c))v)\,dc \\ &= q^m \int_{1+\mathfrak{p}^m} p(\pi(d(c))v)\,dc \\ &= q^m \int_{1+\mathfrak{p}^m} p(v)\,dc \\ &= w. \end{aligned}$$

Evidently, v_1 has property iv). To see properties ii) and iii), let x be in $\mathfrak{p}^{-1}$. By ii) of Lemma 7.2.1,

$$\pi(z(x))v_1 = q^m \int_{\mathfrak{o}^\times} \pi(d(c)z(xc))v\,dc = q^m \int_{\mathfrak{o}^\times} \psi(xc)\pi(d(c))v\,dc.$$

It follows that v_1 has property ii). Integrating over x in $\mathfrak{p}^{-1}$ shows that v_1 has property iii) as well.

To prove that v_1 is unique, let V_1 be the subspace of V consisting of vectors v satisfying properties ii), iii) and iv). We will prove that the restriction of p to V_1 is injective (so that p induces an isomorphism $V_1 \cong V_{Z^J,\psi}$). Let v be in V_1 and assume that $p(v) = 0$. Then there exists a positive integer m such that

$$\int\limits_{\mathfrak{p}^{-m}} \psi(-x)\pi(z(x))v\,dx = 0.$$

Applying $d(c)$ to this equation, where c is in $\mathfrak{o}^\times$, leads to

$$\int\limits_{\mathfrak{p}^{-m}} \psi(-cx)\pi(z(x))v\,dx = 0.$$

Integrating over c in $\mathfrak{o}^\times$, we obtain

$$q^{-1}\int\limits_{\mathfrak{p}^{-1}} \pi(z(x))v\,dx = \int\limits_{\mathfrak{o}} \pi(z(x))v\,dx.$$

Using properties ii) and iii) it follows that $v = 0$. This concludes the proof. ■

7.2.4 Corollary. *Let (π, V) be an irreducible, admissible representation of* $\mathrm{GSp}(4, F)$ *that is not a twist of the trivial representation. Then there exists a vector v in V that is invariant under all elements $d(c)$ with c in $\mathfrak{o}^\times$.*

Proof. By Proposition 7.2.3, it is enough to show that $V_{Z^J,\psi}$ is non-zero. By Proposition 6.1.3, there exists a non-trivial character θ of N such that $V_{N,\theta} \neq 0$. We may assume that θ is of the form (15) with $c = 1$. The assertion follows. ■

References

[1] Jeffrey Adler and Dipendra Prasad. On certain multiplicity one theorems. *Israel J. Math.*, 153:221–245, 2006.

[2] Rolf Berndt and Ralf Schmidt. *Elements of the representation theory of the Jacobi group*, volume 163 of *Progress in Mathematics*. Birkhäuser Verlag, Basel, 1998.

[3] Joseph Bernšteĭn and Andrei Zelevinskiĭ. Representations of the group $GL(n, F)$, where F is a local non-Archimedean field. *Russian Math. Surveys*, 31(3):1–68, 1976.

[4] Daniel Bump, Solomon Friedberg, and Masaaki Furusawa. Explicit formulas for the Waldspurger and Bessel models. *Israel J. Math.*, 102:125–177, 1997.

[5] William Casselman. Introduction to the theory of admissible representations of p-adic reductive groups. Unpublished notes.

[6] Masaaki Furusawa. On L-functions for $\mathrm{GSp}(4) \times \mathrm{GL}(2)$ and their special values. *J. Reine Angew. Math.*, 438:187–218, 1993.

[7] Wee Teck Gan and Atsushi Ichino. On endoscopy and the refined Gross-Prasad conjecture for $(\mathrm{SO}_5, \mathrm{SO}_4)$. *J. Inst. Math. Jussieu*, 10(2):235–324, 2011.

[8] Wee Teck Gan and Shuichiro Takeda. The local Langlands conjecture for GSp(4). *Ann. of Math. (2)*, 173(3):1841–1882, 2011.

[9] Stephen Gelbart and Anthony Knapp. L-indistinguishability and R groups for the special linear group. *Adv. in Math.*, 43(2):101–121, 1982.

[10] Israel Gelfand and David Kazhdan. Representations of the group $\mathrm{GL}(n, K)$ where K is a local field. In *Lie groups and their representations (Proc. Summer School, Bolyai János Math. Soc., Budapest, 1971)*, pages 95–118. Halsted, New York, 1975.

[11] Hervé Jacquet and Robert Langlands. *Automorphic forms on* GL(2), volume 114 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin-New York, 1970.

[12] Stephen Kudla. On the local theta-correspondence. *Invent. Math.*, 83(2):229–255, 1986.

[13] Stephen Kudla and Stephen Rallis. A regularized Siegel-Weil formula: the first term identity. *Ann. of Math. (2)*, 140(1):1–80, 1994.

[14] Colette Mœglin, Marie-France Vignéras, and Jean-Loup Waldspurger. *Correspondances de Howe sur un corps p-adique*, volume 1291 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1987.

[15] Mark Novodvorsky. Uniqueness theorems for generalized Bessel models. *Mat. Sb. (N.S.)*, 90(132):275–287, 326, 1973.

[16] Mark Novodvorsky. Automorphic L-functions for symplectic group GSp(4). In *Automorphic forms, representations and L-functions (Proc. Sympos. Pure Math., Oregon State Univ., Corvallis, Ore., 1977), Part 2*, Proc. Sympos. Pure Math., XXXIII, pages 87–95. Amer. Math. Soc., Providence, R.I., 1979.

[17] Mark Novodvorsky and Ilya Piatetski-Shapiro. Generalized Bessel models for the symplectic group of rank 2. *Mat. Sb. (N.S.)*, 90(132):246–256, 326, 1973.

[18] Timothy O'Meara. *Introduction to quadratic forms*. Classics in Mathematics. Springer-Verlag, Berlin, 2000. Reprint of the 1973 edition.

[19] Ilya Piatetski-Shapiro. L-functions for GSp_4. *Pacific J. Math.*, (Special Issue):259–275, 1997. Olga Taussky-Todd: in memoriam.

[20] Ameya Pitale. Steinberg representation of GSp(4): Bessel models and integral representation of L-functions. *Pacific J. Math.*, 250(2):365–406, 2011.

[21] Ameya Pitale, Abhishek Saha, and Ralf Schmidt. Transfer of Siegel cusp forms of degree 2. *Mem. Amer. Math. Soc.*, 232(1090), 2014.

[22] Ameya Pitale and Ralf Schmidt. Bessel models for GSp(4): Siegel vectors of square-free level. *J. Number Theory*, 136:134–164, 2014.

[23] Dipendra Prasad and Ramin Takloo-Bighash. Bessel models for GSp(4). *J. Reine Angew. Math.*, 655:189–243, 2011.

[24] Brooks Roberts. The theta correspondence for similitudes. *Israel J. Math.*, 94:285–317, 1996.

[25] Brooks Roberts. Tempered representations and the theta correspondence. *Canad. J. Math.*, 50(5):1105–1118, 1998.

[26] Brooks Roberts. The non-Archimedean theta correspondence for GSp(2) and GO(4). *Trans. Amer. Math. Soc.*, 351(2):781–811, 1999.

[27] Brooks Roberts. Global L-packets for GSp(2) and theta lifts. *Doc. Math.*, 6:247–314 (electronic), 2001.

[28] Brooks Roberts and Ralf Schmidt. *Local newforms for GSp(4)*, volume 1918 of *Lecture Notes in Mathematics*. Springer, Berlin, 2007.

[29] Brooks Roberts and Ralf Schmidt. On the number of local newforms in a metaplectic representation. In *Arithmetic geometry and automorphic forms*, volume 19 of *Adv. Lect. Math. (ALM)*, pages 505–530. Int. Press, Somerville, MA, 2011.

[30] François Rodier. Whittaker models for admissible representations of reductive p-adic split groups. In *Harmonic analysis on homogeneous spaces (Proc. Sympos. Pure Math., Vol. XXVI, Williams Coll., Williamstown, Mass., 1972)*, pages 425–430. Amer. Math. Soc., Providence, R.I., 1973.

[31] François Rodier. Les représentations de $\mathrm{GSp}(4,k)$ où k est un corps local. *C. R. Acad. Sci. Paris Sér. A-B*, 283(7):Ai, A429–A431, 1976.

[32] Abhishek Saha. L-functions for holomorphic forms on $\mathrm{GSp}(4)\times\mathrm{GL}(2)$ and their special values. *Int. Math. Res. Not. IMRN*, (10):1773–1837, 2009.

[33] Abhishek Saha. On ratios of Petersson norms for Yoshida lifts. *Forum Math.*, 27(4):2361–2412, 2015.

[34] Hiroshi Saito. On Tunnell's formula for characters of GL(2). *Compositio Math.*, 85(1):99–108, 1993.

[35] Paul Sally, Jr. and Marko Tadić. Induced representations and classifications for $\mathrm{GSp}(2,F)$ and $\mathrm{Sp}(2,F)$. *Mém. Soc. Math. France (N.S.)*, (52):75–133, 1993.

[36] Takashi Sugano. On holomorphic cusp forms on quaternion unitary groups of degree 2. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.*, 31(3):521–568, 1984.

[37] Ramin Takloo-Bighash. L-functions for the p-adic group GSp(4). *Amer. J. Math.*, 122(6):1085–1120, 2000.

[38] Jerrold Tunnell. Local ϵ-factors and characters of GL(2). *Amer. J. Math.*, 105(6):1277–1307, 1983.

[39] Jean-Loup Waldspurger. Correspondance de Shimura. *J. Math. Pures Appl. (9)*, 59(1):1–132, 1980.

[40] Jean-Loup Waldspurger. Sur les valeurs de certaines fonctions L automorphes en leur centre de symétrie. *Compositio Math.*, 54(2):173–242, 1985.

[41] Jean-Loup Waldspurger. Correspondances de Shimura et quaternions. *Forum Math.*, 3(3):219–307, 1991.

Brooks Roberts
Department of Mathematics
University of Idaho
Moscow, ID 83844-1103
USA
brooksr@uidaho.edu

Ralf Schmidt
Department of Mathematics
University of Oklahoma
Norman, OK 73019-3103
USA
rschmidt@math.ou.edu

On Zeta Elements for $\mathbb{G}_m$

David Burns, Masato Kurihara, and Takamichi Sano

Received: November 18, 2015

Communicated by Otmar Venjakob

Abstract. In this paper, we present a unifying approach to the general theory of abelian Stark conjectures. To do so we define natural notions of 'zeta element', of 'Weil-étale cohomology complexes' and of 'integral Selmer groups' for the multiplicative group $\mathbb{G}_m$ over finite abelian extensions of number fields. We then conjecture a precise connection between zeta elements and Weil-étale cohomology complexes, we show this conjecture is equivalent to a special case of the equivariant Tamagawa number conjecture and we give an unconditional proof of the analogous statement for global function fields. We also show that the conjecture entails much detailed information about the arithmetic properties of generalized Stark elements including a new family of integral congruence relations between Rubin-Stark elements (that refines recent conjectures of Mazur and Rubin and of the third author) and explicit formulas in terms of these elements for the higher Fitting ideals of the integral Selmer groups of $\mathbb{G}_m$, thereby obtaining a clear and very general approach to the theory of abelian Stark conjectures. As first applications of this approach, we derive, amongst other things, a proof of (a refinement of) a conjecture of Darmon concerning cyclotomic units, a proof of (a refinement of) Gross's 'Conjecture for Tori' in the case that the base field is $\mathbb{Q}$, explicit conjectural formulas for both annihilating elements and, in certain cases, the higher Fitting ideals (and hence explicit structures) of ideal class groups and a strong refinement of many previous results concerning abelian Stark conjectures.

Contents

The second and the third authors are partially supported by JSPS Core-to-core program, "Foundation of a Global Research Cooperative Center in Mathematics focused on Number Theory and Geometry".

1. Introduction

The study of the special values of zeta functions and, more generally, of L-functions is a central theme in number theory that has a long tradition stretching back to Dirichlet and Kummer in the nineteenth century. In particular, much work has been done concerning the arithmetic properties of the special values of L-functions and their incarnations in appropriate arithmetic cohomology groups, or 'zeta elements' as they are commonly known.

The aim of our project is to systematically study the fine arithmetic properties of such zeta elements and thereby to obtain both generalizations and refinements of a wide range of well-known results and conjectures in the area. In this first article we shall concentrate, for primarily pedagogical reasons, on the classical and very concrete case of the L-functions that are attached to the multiplicative group $\mathbb{G}_m$ over a finite abelian extension K/k of global fields. In subsequent articles we will then investigate the key Iwasawa-theoretic aspects of our approach (see [9]) and also explain how the conjectures and results presented here naturally extend both to the case of Galois extensions that are not abelian and to the case of the zeta elements that are associated (in general conjecturally) to a wide class of motives over number fields.

The main results of the present article are given below as Theorems 1.1, 1.5 and 1.10. In the rest of this introduction we state these results and also discuss a selection of interesting consequences.

To do this we fix a finite abelian extension of global fields K/k with Galois group $G = \mathrm{Gal}(K/k)$.

We then fix a finite non-empty set of places S of k containing both the set $S_{\mathrm{ram}}(K/k)$ of places which ramify in K/k and the set $S_\infty(k)$ of archimedean places (if any).

Lastly we fix an auxiliary finite non-empty set of places T of k which is disjoint from S and such that the group $\mathcal{O}_{K,S,T}^\times$ of S-units of K that are endowed with a trivialization at each place of K above a place in T is $\mathbb{Z}$-torsion-free (for the precise definition of $\mathcal{O}_{K,S,T}^\times$, see §1.7).

1.1. The leading term conjecture and Rubin-Stark elements.

As a first step we shall define a canonical 'T-modified Weil-étale cohomology' complex for $\mathbb{G}_m$ and then formulate (as Conjecture 3.6) a precise 'leading term conjecture' $\mathrm{LTC}(K/k)$ for the extension K/k. This conjecture predicts that the canonical *zeta element* $z_{K/k,S,T}$ interpolating the leading terms at $s = 0$ of the (S-truncated T-modified) L-functions $L_{k,S,T}(\chi, s)$ generates the determinant module over G of the T-modified Weil-étale cohomology complex for $\mathbb{G}_m$ over K.

The main result of the first author in [5] implies that $\mathrm{LTC}(K/k)$ is valid if k is a global function field.

In the number field case our formulation of $\mathrm{LTC}(K/k)$ is motivated by the 'Tamagawa Number Conjecture' formulated by Bloch and Kato in [1] and by the 'generalized Iwasawa main conjecture' studied by Kato in [24] and [25]. In particular, we shall show that for extensions K/k of number fields $\mathrm{LTC}(K/k)$

is equivalent to the relevant special case of the 'equivariant Tamagawa number conjecture' formulated in the article [7] of Flach and the first author. Taken in conjunction with previous work of several authors, this fact implies that $\mathrm{LTC}(K/k)$ is also unconditionally valid for several important families of number fields.

We assume now that S contains a subset $V = \{v_1, \ldots, v_r\}$ of places which split completely in K. In this context, one can use the values at $s = 0$ of the r-th derivatives of S-truncated T-modified L-functions to define a canonical element

$$\epsilon^V_{K/k,S,T}$$

in the exterior power module $\bigwedge^r_{\mathbb{Z}[G]} \mathcal{O}^\times_{K,S,T} \otimes \mathbb{R}$ (for the precise definition see §5.1).

As a natural generalization of a classical conjecture of Stark (dealing with the case $r = 1$) Rubin conjectured in [45] that the elements $\epsilon^V_{K/k,S,T}$ should always satisfy certain precise integrality conditions (for more details see Remark 1.6). As is now common in the literature, in the sequel we shall refer to $\epsilon^V_{K/k,S,T}$ as the 'Rubin-Stark element' (relative to the given data) and to the central conjecture of Rubin in [45] as the 'Rubin-Stark Conjecture'.

In some very special cases $\epsilon^V_{K/k,S,T}$ can be explicitly computed and the Rubin-Stark Conjecture verified. For example, this is the case if $r = 0$ (so $V = \emptyset$) when $\epsilon^V_{K/k,S,T}$ can be described in terms of Stickelberger elements and if $k = \mathbb{Q}$ and $V = \{\infty\}$ when $\epsilon^V_{K/k,S,T}$ can be described in terms of cyclotomic units.

As a key step in our approach we show that in all cases the validity of $\mathrm{LTC}(K/k)$ implies that $\epsilon^V_{K/k,S,T}$ can be computed as 'the canonical projection' of the zeta element $z_{K/k,S,T}$.

This precise result is stated as Theorem 5.14 and its proof will also incidentally show that $\mathrm{LTC}(K/k)$ implies the validity of the Rubin-Stark conjecture for K/k. The latter implication was in fact already observed by the first author in [3] (and the techniques developed in loc. cit. have since been used by several other authors) but we would like to point out that the proof presented here is very much simpler than that given in [3] and is therefore much more amenable to subsequent generalization.

1.2. Refined class number formulas for $\mathbb{G}_m$. The first consequence of Theorem 5.14 that we record here concerns a refined version of a conjecture that was recently formulated independently by Mazur and Rubin in [37] (where it is referred to as a 'refined class number formula for $\mathbb{G}_m$') and by the third author in [46].

To discuss this we fix an intermediate field L of K/k and a subset $V' = \{v_1, \ldots, v_{r'}\}$ of S which contains V and is such that every place in V' splits completely in L.

In this context it is known that the elements $\epsilon^V_{K/k,S,T}$ naturally constitute an Euler system of rank r and the elements $\epsilon^{V'}_{L/k,S,T}$ an Euler system of rank r'. If $r < r'$, then the image of $\epsilon^V_{K/k,S,T}$ under the map induced by the field theoretic

norm $K^\times \to L^\times$ vanishes. However, in this case Mazur and Rubin (see [37, Conj. 5.2]) and the third author (see [46, Conj. 3]) independently observed that the reciprocity maps of local class field theory lead to an important conjectural relationship between the elements $\epsilon^V_{K/k,S,T}$ and $\epsilon^{V'}_{L/k,S,T}$.
We shall here formulate an interesting refinement $\mathrm{MRS}(K/L/k,S,T)$ of the central conjectures of [37] and [46] (see Conjecture 5.4 and the discussion of Remark 5.7) and we shall then prove the following result.

Theorem 1.1. $\mathrm{LTC}(K/k)$ *implies the validity of* $\mathrm{MRS}(K/L/k,S,T)$.

This result is both a generalization and strengthening of the main result of the third author in [46, Th. 3.22] and provides strong evidence for $\mathrm{MRS}(K/L/k,S,T)$.
As already remarked earlier, if k is a global function field, then the validity of $\mathrm{LTC}(K/k)$ is a consequence of the main result of [5]. In addition, if $k=\mathbb{Q}$, then the validity of $\mathrm{LTC}(K/k)$ follows from the work of Greither and the first author in [8] and of Flach in [14].
Theorem 1.1 therefore has the following consequence.

Corollary 1.2. $\mathrm{MRS}(K/L/k,S,T)$ *is valid if* $k=\mathbb{Q}$ *or if* k *is a global function field.*

This result is of particular interest since it verifies the conjectures of Mazur and Rubin [37] and of the third author [46] even in cases for which one has $r>1$.
In a sequel [9] to this article we will also prove a partial converse to Theorem 1.1 and show that this converse can be used to derive significant new evidence in support of the conjecture $\mathrm{LTC}(K/k)$ (for more details see §1.6 below).
Next we recall that in [12] Darmon used the theory of cyclotomic units to formulate a refined version of the class number formula for the class groups of real quadratic fields. We further recall that Mazur and Rubin in [36], and later the third author in [46], have proved the validity of the central conjecture of [12] but only after inverting the prime 2.
We shall formulate in §6 a natural refinement of Darmon's conjecture. By using Corollary 1.2 we shall then give a full proof of our refined version of Darmon's conjecture, thereby obtaining the following result (for a precise version of which see Theorem 6.1).

Corollary 1.3. *A natural refinement of Darmon's conjecture in* [12] *is valid.*

Let now K/k be an abelian extension as above and choose intermediate fields L and $\widetilde{L}$ with $[L:k]=2$, $L\cap\widetilde{L}=k$ and $K=L\widetilde{L}$. In this context Gross has formulated in [21] a 'conjecture for tori' regarding the value of the canonical Stickelberger element associated to K/k modulo a certain ideal constructed from class numbers and a canonical integral regulator map. This conjecture has been widely studied in the literature, perhaps most notably by Hayward in [22] and by Greither and Kučera in [16, 17].

We shall formulate (as Conjecture 6.3) a natural refinement of Gross's conjecture for tori and we shall then prove (in Theorem 6.5) that the validity of this refinement is a consequence of $\mathrm{MRS}(K/L/k, S, T)$.

As a consequence of Corollary 1.2 we shall therefore obtain the following result.

Corollary 1.4. *A natural refinement of Gross's conjecture for tori is valid if $k = \mathbb{Q}$ or if k is a global function field.*

This result is a significant improvement of the main results of Greither and Kučera in [16, 17]. In particular, whilst the latter articles only study the case that $k = \mathbb{Q}$, L is an imaginary quadratic field, and $\widetilde{L}/\mathbb{Q}$ is an abelian extension satisfying several technical conditions (see Remark 6.6), Corollary 1.4 now proves Gross's conjecture completely in the case $k = \mathbb{Q}$ and with no assumption on either L or $\widetilde{L}$.

1.3. Selmer groups and their higher Fitting ideals. In order to state our second main result, we introduce two new Galois modules which are each finitely generated abelian groups and will play a key role in the arithmetic theory of zeta elements.

The first of these is a canonical '(Σ-truncated T-modified) integral dual Selmer group' $\mathcal{S}_{\Sigma,T}(\mathbb{G}_{m/K})$ for the multiplicative group over K for each finite non-empty set of places Σ of K that contains $S_\infty(K)$ and each finite set of places T of K that is disjoint from Σ.

If $\Sigma = S_\infty(K)$ and T is empty, then $\mathcal{S}_{\Sigma,T}(\mathbb{G}_{m/K})$ is simply defined to be the cokernel of the map

$$\prod_w \mathbb{Z} \longrightarrow \mathrm{Hom}_{\mathbb{Z}}(K^\times, \mathbb{Z}), \quad (x_w)_w \mapsto (a \mapsto \sum_w \mathrm{ord}_w(a) x_w),$$

where in the product and sum w runs over all finite places of K, and in this case constitutes a canonical integral structure on the Pontryagin dual of the Bloch-Kato Selmer group $H^1_f(K, \mathbb{Q}/\mathbb{Z}(1))$ (see Remark 2.3(i)).

In general, the group $\mathcal{S}_{\Sigma,T}(\mathbb{G}_{m/K})$ is defined to be a natural analogue for $\mathbb{G}_m$ of the 'integral Selmer group' that was introduced for abelian varieties by Mazur and Tate in [38] and, in particular, lies in a canonical exact sequence of G-modules of the form

(1)
$$0 \longrightarrow \mathrm{Hom}_{\mathbb{Z}}(\mathrm{Cl}^T_\Sigma(K), \mathbb{Q}/\mathbb{Z}) \longrightarrow \mathcal{S}_{\Sigma,T}(\mathbb{G}_{m/K}) \longrightarrow \mathrm{Hom}_{\mathbb{Z}}(\mathcal{O}^\times_{K,\Sigma,T}, \mathbb{Z}) \longrightarrow 0$$

where $\mathrm{Cl}^T_\Sigma(K)$ is the ray class group of $\mathcal{O}_{K,\Sigma}$ modulo the product of all places of K above T (see §1.7).

This Selmer group is also philosophically related to the theory of Weil-étale cohomology that is conjectured to exist by Lichtenbaum in [34], and in this direction we show that in all cases there is a natural identification

$$\mathcal{S}_{\Sigma,T}(\mathbb{G}_{m/K}) = H^2_{c,T}((\mathcal{O}_{K,\Sigma})_{\mathcal{W}}, \mathbb{Z})$$

where the right hand group denotes the cohomology in degree two of a canonical 'T-modified compactly supported Weil-étale cohomology complex' that we introduce in §2.2.

The second module $\mathcal{S}^{\rm tr}_{\Sigma,T}(\mathbb{G}_{m/K})$ that we introduce is a canonical 'transpose' (in the sense of Jannsen's homotopy theory of modules [23]) for $\mathcal{S}_{\Sigma,T}(\mathbb{G}_{m/K})$. In terms of the complexes introduced in §2.2 this module can be described as a certain 'T-modified Weil-étale cohomology group' of $\mathbb{G}_m$

$$\mathcal{S}^{\rm tr}_{\Sigma,T}(\mathbb{G}_{m/K}) = H^1_T((\mathcal{O}_{K,\Sigma})_{\mathcal{W}}, \mathbb{G}_m)$$

and can also be shown to lie in a canonical exact sequence of G-modules of the form

$$0 \longrightarrow {\rm Cl}^T_\Sigma(K) \longrightarrow \mathcal{S}^{\rm tr}_{\Sigma,T}(\mathbb{G}_{m/K}) \longrightarrow X_{K,\Sigma} \longrightarrow 0. \tag{2}$$

Here $X_{K,\Sigma}$ denotes the subgroup of the free abelian group on the set Σ_K of places of K above Σ comprising elements whose coefficients sum to zero.
We can now state our second main result.
In this result we write ${\rm Fitt}^r_G(M)$ for the r-th Fitting ideal of a finitely generated G-module M, though the usual notation is ${\rm Fitt}_{r,\mathbb{Z}[G]}(M)$, in order to make the notation consistent with the exterior power $\bigwedge^r_{\mathbb{Z}[G]} M$. Note that we will review the definition of higher Fitting ideals in §7.1 and also introduce there for each finitely generated G-module M and each pair of non-negative integers r and i a natural notion of 'higher relative Fitting ideal'

$${\rm Fitt}^{(r,i)}_G(M) = {\rm Fitt}^{(r,i)}_G(M, M_{\rm tors}).$$

We write $x \mapsto x^\#$ for the $\mathbb{C}$-linear involution of $\mathbb{C}[G]$ which inverts elements of G.

Theorem 1.5. *Let $K/k, S, T, V$ and r be as above, and assume that ${\rm LTC}(K/k)$ is valid. Then all of the following claims are also valid.*

(i) *One has*

$${\rm Fitt}^r_G(\mathcal{S}_{S,T}(\mathbb{G}_{m/K})) = \{\Phi(\epsilon^V_{K/k,S,T})^\# : \Phi \in \bigwedge_{\mathbb{Z}[G]}^r {\rm Hom}_{\mathbb{Z}[G]}(\mathcal{O}^\times_{K,S,T}, \mathbb{Z}[G])\}.$$

(ii) *Let $\mathcal{P}_k(K)$ be the set of all places which split completely in K. Fix a non-negative integer i and set*

$$\mathcal{V}_i = \{V' \subset \mathcal{P}_k(K) : |V'| = i \text{ and } V' \cap (S \cup T) = \emptyset\}.$$

Then one has

$$\begin{aligned}&{\rm Fitt}^{(r,i)}_G(\mathcal{S}^{\rm tr}_{S,T}(\mathbb{G}_{m/K}))\\ = \ &\{\Phi(\epsilon^{V\cup V'}_{K/k,S\cup V',T}) : V' \in \mathcal{V}_i \text{ and } \Phi \in \bigwedge_{\mathbb{Z}[G]}^{r+i} {\rm Hom}_{\mathbb{Z}[G]}(\mathcal{O}^\times_{K,S\cup V',T}, \mathbb{Z}[G])\}.\end{aligned}$$

In particular, if $i = 0$, then one has

$${\rm Fitt}^r_G(\mathcal{S}^{\rm tr}_{S,T}(\mathbb{G}_{m/K})) = \{\Phi(\epsilon^V_{K/k,S,T}) : \Phi \in \bigwedge_{\mathbb{Z}[G]}^r {\rm Hom}_{\mathbb{Z}[G]}(\mathcal{O}^\times_{K,S,T}, \mathbb{Z}[G])\}.$$

Remark 1.6. In terms of the notation of Theorem 1.5, the Rubin-Stark Conjecture asserts that $\Phi(\epsilon^V_{K/k,S,T})$ belongs to $\mathbb{Z}[G]$ for every Φ in $\bigwedge^r_{\mathbb{Z}[G]}\mathrm{Hom}_{\mathbb{Z}[G]}(\mathcal{O}^\times_{K,S,T},\mathbb{Z}[G])$. The property described in Theorem 1.5 is deeper in that it shows the ideal generated by $\Phi(\epsilon^V_{K/k,S,T})$ as Φ runs over $\bigwedge^r_{\mathbb{Z}[G]}\mathrm{Hom}_{\mathbb{Z}[G]}(\mathcal{O}^\times_{K,S,T},\mathbb{Z}[G])$ should encode significant arithmetic information relating to integral Selmer groups. (See also Remark 5.13 in this regard.)

1.4. Galois structures of ideal class groups. In this subsection, in order to better understand the content of Theorem 1.5, we discuss several interesting consequences concerning the explicit Galois structure of ideal class groups.

To do this we fix an odd prime p and suppose that K/k is any finite abelian extension of global fields. We write L for the (unique) intermediate field of K/k such that K/L is a p-extension and $[L:k]$ is prime to p. Then the group $\mathrm{Gal}(K/k)$ decomposes as a direct product $\mathrm{Gal}(L/k)\times\mathrm{Gal}(K/L)$ and we fix a non-trivial faithful character χ of $\mathrm{Gal}(L/k)$. We set $\mathrm{Cl}^T(K) := \mathrm{Cl}^T_\emptyset(K)$ and define its '(p,χ)-component' by setting

$$A^T(K)^\chi := (\mathrm{Cl}^T(K)\otimes\mathbb{Z}_p)\otimes_{\mathbb{Z}_p[\mathrm{Gal}(L/k)]}\mathcal{O}_\chi.$$

Here we write $\mathcal{O}_\chi$ for the module $\mathbb{Z}_p[\mathrm{im}(\chi)]$ upon which $\mathrm{Gal}(L/k)$ acts via χ so that $A^T(K)^\chi$ has an induced action of the group ring $R^\chi_K := \mathcal{O}_\chi[\mathrm{Gal}(K/L)]$.

Then in Theorem 8.1 we shall derive the following results about the structure of $A^T(K)^\chi$ from the final assertion of Theorem 1.5(ii).

In this result we write '$\chi(v)\neq 1$' to mean that the decomposition group of v in $\mathrm{Gal}(L/k)$ is non-trivial.

Corollary 1.7. *Let V be the set of archimedean places of k that split completely in K and set $r := |V|$. Assume that any ramifying place v of k in K satisfies $\chi(v)\neq 1$. Assume also that the equality of $\mathrm{LTC}(K/k)$ is valid after applying the functor $-\otimes_{\mathbb{Z}_p[\mathrm{Gal}(L/k)]}\mathcal{O}_\chi$.*

Then for any non-negative integer i one has an equality

$$\mathrm{Fitt}^i_{R^\chi_K}(A^T(K)^\chi) = \{\Phi(\epsilon^{V\cup V',\chi}_{K/k,S\cup V',T}) : V'\in\mathcal{V}_i \text{ and } \Phi\in\bigwedge^{r+i}_{R^\chi_K}\mathcal{H}_\chi\}$$

where we set $S := S_\infty(k)\cup S_{\mathrm{ram}}(K/k)$ and $\mathcal{H}_\chi := \mathrm{Hom}_{R^\chi_K}((\mathcal{O}^\times_{K,S\cup V',T}\otimes\mathbb{Z}_p)^\chi, R^\chi_K)$.

We remark that Corollary 1.7 specializes to give refinements of several results in the literature.

For example, if $k=\mathbb{Q}$ and K is equal to the maximal totally real subfield $\mathbb{Q}(\zeta_m)^+$ of $\mathbb{Q}(\zeta_m)$ where ζ_m is a fixed choice of primitive m-th root of unity for some natural number m, then $\mathrm{LTC}(K/k)$ is known to be valid and so Corollary 1.7 gives an explicit description of the higher Fitting ideals of ideal class groups in terms of cyclotomic units (which are the relevant Rubin-Stark elements in this case). In particular, if $m=p^n$ for any non-negative integer n, then the

necessary condition on χ is satisfied for all non-trivial χ and Corollary 1.7 gives a strong refinement of Ohshita's theorem in [41] for the field $K = \mathbb{Q}(\zeta_{p^n})^+$.
The result is also stronger than that of Mazur and Rubin in [35, Th. 4.5.9] since the latter describes structures over a discrete valuation ring whilst Corollary 1.7 describes structures over the group ring R_K^χ.
In addition, if K is a CM extension of a totally real field k, then Corollary 1.7 constitutes a generalization of the main results of the second author in both [28] and [30]. To explain this we suppose that K/k is a CM-extension and that χ is an odd character. Then classical Stickelberger elements can be used to define for each non-negative integer i a 'higher Stickelberger ideal'

$$\Theta^i(K/k) \subseteq \mathbb{Z}_p[\mathrm{Gal}(K/k)]$$

(for details see §8.3). By taking T to be empty we can consider the (p,χ)-component of the usual ideal class group

$$A(K)^\chi := (\mathrm{Cl}^T(K) \otimes \mathbb{Z}_p) \otimes_{\mathbb{Z}_p[\mathrm{Gal}(L/k)]} \mathcal{O}_\chi.$$

Then, by using both Theorem 1.1 and Corollary 1.7 we shall derive the following result as a consequence of the more general Theorem 8.6.
In this result we write ω for the Teichmüller character giving the Galois action on the group of p-th roots of unity.

Corollary 1.8. *Let K be a CM-field, k totally real, and χ an odd character with $\chi \neq \omega$. We assume that any ramifying place v of k in K satisfies $\chi(v) \neq 1$ and that $\mathrm{LTC}(F/k)$ is valid for certain extensions F of K (see Theorem 8.6 for the precise conditions on F).*
Then for any non-negative integer i one has an equality

$$\mathrm{Fitt}^i_{R_K^\chi}(A(K)^\chi) = \Theta^i(K/k)^\chi.$$

In the notation of Corollary 1.8 suppose that K is the n-th layer of the cyclotomic $\mathbb{Z}_p$-extension of L for some non-negative integer n and that every place $\mathfrak{p}$ above p satisfies $\chi(\mathfrak{p}) \neq 1$. Then the conditions on $\chi(v)$ and $\mathrm{LTC}(F/k)$ that are stated in Corollary 1.8 are automatically satisfied and so Corollary 1.8 generalizes the main results of the second author in [30].
To get a better feeling for Corollary 1.8, consider the simple case that $[K:k]$ is prime to p. In this case $K = L$, the ring $\mathbb{Z}_p[\mathrm{Gal}(K/k)]$ is semi-local and $A(K)^\chi$ is a module over the discrete valuation ring $\mathcal{O}_\chi = R_K^\chi$. Then the conclusion in Corollary 1.8 in the case $i = 0$ implies that

$$|A(K)^\chi| = |\mathcal{O}_\chi / L_k(\chi^{-1}, 0)| \tag{3}$$

where $L_k(\chi^{-1}, s)$ is the usual Artin L-function. If every place $\mathfrak{p}$ above p satisfies $\chi(\mathfrak{p}) \neq 1$, then this equality is known to be a consequence of the main conjecture for totally real fields proved by Wiles [54]. However, without any such restriction on the values $\chi(\mathfrak{p})$, the equality (3) is as yet unproved.
In addition, in this case the result of Corollary 1.8 is much stronger than (3) in that it shows the explicit structure of $A(K)^\chi$ as a Galois module to be

completely determined (conjecturally at least) by Stickelberger elements by using the obvious (non-canonical) isomorphism of $\mathcal{O}_\chi$-modules

$$A(K)^\chi \simeq \bigoplus_{i\geq 1} \mathrm{Fitt}^i_{\mathcal{O}_\chi}(A(K)^\chi)/\,\mathrm{Fitt}^{i-1}_{\mathcal{O}_\chi}(A(K)^\chi) = \bigoplus_{i\geq 1} \Theta^i(K/k)^\chi/\Theta^{i-1}(K/k)^\chi.$$

Next we note that the proof of Corollary 1.8 also combines with the result of Theorem 1.16 below to give the following result (which does not itself assume the validity of $\mathrm{LTC}(K/k)$).

This result will be proved in Corollaries 8.4 and 8.8. In it we write $\mu_{p^\infty}(k(\zeta_p))$ for the p-torsion subgroup of $k(\zeta_p)^\times$.

Corollary 1.9. *Assume that K/k is a CM-extension, that the degree $[K:k]$ is prime to p, and that χ is an odd character of G such that there is at most one p-adic place $\mathfrak{p}$ of k with $\chi(\mathfrak{p}) = 1$. Assume also that the p-adic μ-invariant of K_∞/K vanishes.*

Then one has both an equality

$$|A(K)^\chi| = \begin{cases} |\mathcal{O}_\chi/L_k(\chi^{-1},0)| & \text{if } \chi \neq \omega, \\ |\mathcal{O}_\chi/(|\mu_{p^\infty}(k(\zeta_p))|\cdot L_k(\chi^{-1},0))| & \text{if } \chi = \omega \end{cases}$$

and a (non-canonical) isomorphism of $\mathcal{O}_\chi$-modules

$$A(K)^\chi \simeq \bigoplus_{i\geq 1} \Theta^i(K/k)^\chi/\Theta^{i-1}(K/k)^\chi.$$

This result is a generalization of the main theorem of the second author in [28] where it is assumed that $\chi(\mathfrak{p}) \neq 1$ for all p-adic places $\mathfrak{p}$. It also generalizes the main result of Rubin in [44] which deals only with the special case $K = \mathbb{Q}(\mu_p)$ and $k = \mathbb{Q}$.

To end this subsection we note Remark 1.13 below explains why Theorem 1.5(ii) also generalizes and refines the main result of Cornacchia and Greither in [10].

1.5. Annihilators and Fitting ideals of class groups for small Σ.

In this subsection we discuss further connections between Rubin-Stark elements and the structure of class groups of the form $\mathrm{Cl}^T_\Sigma(K)$ for 'small' sets Σ which do not follow from Theorem 1.5. In particular, we do not assume here that Σ contains $S_{\mathrm{ram}}(K/k)$.

To do this we denote the annihilator ideal of a G-module M by $\mathrm{Ann}_G(M)$.

Theorem 1.10. *Assume $\mathrm{LTC}(K/k)$ is valid.*

Fix Φ in $\bigwedge^r_{\mathbb{Z}[G]}\mathrm{Hom}_{\mathbb{Z}[G]}(\mathcal{O}^\times_{K,S,T},\mathbb{Z}[G])$ and any place v in $S\setminus V$.

Then one has

$$\Phi(\epsilon^V_{K/k,S,T}) \in \mathrm{Ann}_G(\mathrm{Cl}^T_{V\cup\{v\}}(K))$$

and, if G is cyclic, also

$$\Phi(\epsilon^V_{K/k,S,T}) \in \mathrm{Fitt}^0_G(\mathrm{Cl}^T_{V\cup\{v\}}(K)).$$

Remark 1.11. The first assertion of Theorem 1.10 provides a common refinement and wide-ranging generalization (to L-series of arbitrary order of vanishing) of several well-known conjectures and results in the literature. To discuss this we write $\mathrm{Cl}^T(K)$ for the full ray class group modulo T (namely, $\mathrm{Cl}^T(K) = \mathrm{Cl}^T_\emptyset(K)$, see §1.7).
(i) We first assume that $r = 0$ (so V is empty) and that k is a number field. Then, without loss of generality (for our purposes), we can assume that k is totally real and K is a CM field. In this case $\epsilon^{\emptyset}_{K/k,S,T}$ is the Stickelberger element $\theta_{K/k,S,T}(0)$ of the extension K/k (see §3.1). We take v to be an archimedean place in S. Then $\mathrm{Cl}^T_{\{v\}}(K) = \mathrm{Cl}^T(K)$ and so the first assertion of Theorem 1.10 shows that $\mathrm{LTC}(K/k)$ implies the classical Brumer-Stark Conjecture,

$$\theta_{K/k,S,T}(0) \cdot \mathrm{Cl}^T(K) = 0.$$

(ii) We next consider the case that K is totally real and take V to be $S_\infty(k)$ so that $r = |V| = [k : \mathbb{Q}]$. In this case Corollary 1.10 implies that for any non-archimedean place v in S, any element σ_v of the decomposition subgroup G_v of v in G and any element Φ of $\bigwedge^{[k:\mathbb{Q}]}_{\mathbb{Z}[G]} \mathrm{Hom}_{\mathbb{Z}[G]}(\mathcal{O}^\times_{K,S,T}, \mathbb{Z}[G])$, one has

$$(1 - \sigma_v) \cdot \Phi(\epsilon^{S_\infty}_{K/k,S,T}) \in \mathrm{Ann}_G(\mathrm{Cl}^T(K)). \tag{4}$$

To make this containment even more explicit we further specialize to the case that $k = \mathbb{Q}$ and that K is equal to $\mathbb{Q}(\zeta_m)^+$ for some natural number m. We recall that $\mathrm{LTC}(K/k)$ has been verified in this case. We take S to be the set comprising the (unique) archimedean place ∞ and all prime divisors of m, and V to be $S_\infty = \{\infty\}$ (so $r = 1$). For a set T which contains an odd prime, we set $\delta_T := \prod_{v\in T}(1 - \mathrm{N}v\mathrm{Fr}_v^{-1})$, where $\mathrm{Fr}_v \in G$ denotes the Frobenius automorphism at a place of K above v. In this case, we have

$$\epsilon^{\{\infty\}}_{K/k,S,T} = \epsilon_{m,T} := (1 - \zeta_m)^{\delta_T} \in \mathcal{O}^\times_{K,S,T}$$

(see, for example, [50, p.79] or [42, §4.2]) and so (4) implies that for any $\sigma_v \in G_v$ and any $\Phi \in \mathrm{Hom}_{\mathbb{Z}[G]}(\mathcal{O}^\times_{K,S,T}, \mathbb{Z}[G])$ one has

$$(1 - \sigma_v) \cdot \Phi(\epsilon_{m,T}) \in \mathrm{Ann}_G(\mathrm{Cl}^T(K)).$$

Now the group G is generated by the decomposition subgroups G_v of each prime divisor v of m, and so for any $\sigma \in G$ one has an equality $\sigma - 1 = \Sigma_{v|m} x_v$ for suitable elements x_v of the ideal $I(G_v)$ of $\mathbb{Z}[G]$ that is generated by $\{\sigma_v - 1 : \sigma_v \in G_v\}$. Hence, since $\epsilon^{\sigma-1}_{m,T}$ belongs to $\mathcal{O}^\times_K$ one finds that for any $\varphi \in \mathrm{Hom}_{\mathbb{Z}[G]}(\mathcal{O}^\times_K, \mathbb{Z}[G])$ one has $\varphi(\epsilon^{\sigma-1}_{m,T}) = \Sigma_{v|m} x_v \widetilde{\varphi}(\epsilon_{m,T})$ where $\widetilde{\varphi}$ is any lift of φ to $\mathrm{Hom}_{\mathbb{Z}[G]}(\mathcal{O}^\times_{K,S,T}, \mathbb{Z}[G])$. Therefore, for any φ in $\mathrm{Hom}_{\mathbb{Z}[G]}(\mathcal{O}^\times_K, \mathbb{Z}[G])$ and any σ in G, one has

$$\varphi(\epsilon^{\sigma-1}_{m,T}) \in \mathrm{Ann}_G(\mathrm{Cl}^T(K)). \tag{5}$$

This containment is actually finer than the annihilation result proved by Rubin in [43, Th. 2.2 and the following Remark] since it deals with the group $\mathrm{Cl}^T(K)$ rather than $\mathrm{Cl}(K)$.

Remark 1.12. We next consider the case that K/k is a cyclic CM-extension and V is empty. As remarked above, in this case the Rubin-Stark element $\epsilon^{\emptyset}_{K/k,S,T}$ coincides with the Stickelberger element $\theta_{K/k,S,T}(0)$.
The second assertion of Theorem 1.10 therefore combines with the argument in Remark 1.11(i) to show that $\mathrm{LTC}(K/k)$ implies a containment

$$\theta_{K/k,S,T}(0) \in \mathrm{Fitt}^0_G(\mathrm{Cl}^T(K)).$$

This is a strong refinement of the Brumer-Stark conjecture. To see this note that $\mathrm{Cl}^T(K)$ is equal to the ideal class group $\mathrm{Cl}(K)$ of K when T is empty. The above containment thus combines with [50, Chap. IV, Lem. 1.1] to imply that if G is cyclic, then one has

$$\theta_{K/k,S}(0) \cdot \mathrm{Ann}_G(\mu(K)) \subset \mathrm{Fitt}^0_G(\mathrm{Cl}(K))$$

where $\mu(K)$ denotes the group of roots of unity in K. It is known that this inclusion is not in general valid without the hypothesis that G is cyclic (see [18]). The possibility of such an explicit refinement of Brumer's Conjecture was discussed by the second author in [29] and [31]. In fact, in the terminology of [29], the above argument shows that both properties (SB) and (DSB) follow from $\mathrm{LTC}(K/k)$ whenever G is cyclic. For further results in the case that G is cyclic see Corollary 7.10.

Remark 1.13. Following the discussion of Remark 1.11(ii) we can also now consider Theorem 1.5 further in the case that $k = \mathbb{Q}$, $K = \mathbb{Q}(\zeta_{p^n})^+$ for an odd prime p and natural number n and $S = \{\infty, p\}$.
In this case the G-module $X_{K,S}$ is free of rank one and so the exact sequence (2) combines with the final assertion of Theorem 1.5(ii) (with $r = 1$) to give equalities

$$\begin{aligned}\mathrm{Fitt}^0_G(\mathrm{Cl}^T_S(K)) &= \mathrm{Fitt}^1_G(\mathcal{S}^{\mathrm{tr}}_{S,T}(\mathbb{G}_{m/K}))\\ &= \{\Phi(\epsilon_{p^n,T}) \mid \Phi \in \mathrm{Hom}_G(\mathcal{O}^\times_{K,S,T}, \mathbb{Z}[G])\}\\ &= \mathrm{Fitt}^0_G(\mathcal{O}^\times_{K,S,T}/(\mathbb{Z}[G]\cdot \epsilon_{p^n,T}))\end{aligned}$$

where the last equality follows from the fact that G is cyclic.
Since (in this case) $\mathrm{Cl}_S(K) = \mathrm{Cl}(K)$ a standard argument shows that the above displayed equality implies $\mathrm{Fitt}^0_G(\mathrm{Cl}(K)) = \mathrm{Fitt}^0_G(\mathcal{O}^\times_K/\mathcal{C}_K)$ with $\mathcal{C}_K$ denoting the group $\mathbb{Z}[G]\cdot\{1-\zeta_{p^n}, \zeta_{p^n}\} \cap \mathcal{O}^\times_K$ of cyclotomic units of K, and this is the main result of Cornacchia and Greither in [10]. Our results therefore constitute an extension of the main result in [10] for $K = \mathbb{Q}(\zeta_{p^n})^+$.

For any finite group Γ and any Γ-module M we write $M^\vee$ for its Pontryagin dual $\mathrm{Hom}_{\mathbb{Z}}(M, \mathbb{Q}/\mathbb{Z})$, endowed with the natural contragredient action of Γ.
In §7.5 we show that the proof of Theorem 1.10 also implies the following result. In this result we fix an odd prime p and set $\mathrm{Cl}^T(K)^\vee_p := \mathrm{Cl}^T(K)^\vee \otimes \mathbb{Z}_p$.

Corollary 1.14. *Let K/k be any finite abelian CM-extension and p any odd prime. If $\mathrm{LTC}(K/k)$ is valid, then one has a containment*

$$\theta_{K/k,S,T}(0)^\# \in \mathrm{Fitt}^0_{\mathbb{Z}_p[G]}(\mathrm{Cl}^T(K)^\vee_p).$$

Remark 1.15.
(i) In [19] Greither and Popescu prove the validity of the displayed containment in Corollary 1.14 under the hypotheses that S contains all p-adic places of k (so that the Stickelberger element $\theta_{K/k,S,T}(0)$ is in general 'imprimitive') and that the p-adic μ-invariant of K vanishes. In [9, §3.5] we give a new proof of their result by using the natural Selmer modules for $\mathbb{G}_m$ defined in §2 below in place of the Galois modules 'related to 1-motives' that are explicitly constructed for this purpose in [19]. In addition, by combining Corollary 1.14 with the result of Theorem 1.16 below we can also now prove the containment in Corollary 1.14, both unconditionally and without the assumption that S contains all p-adic places, for important families of examples. For more details see [9, §3.5 and §5].
(ii) For any odd prime p the group $\mathrm{Cl}(K)_p^\vee := \mathrm{Cl}(K)^\vee \otimes \mathbb{Z}_p$ is not a quotient of $\mathrm{Cl}^T(K)_p^\vee$ and so Corollary 1.14 does not imply that $\theta_{K/k,S,T}(0)$ belongs to $\mathrm{Fitt}^0_{\mathbb{Z}_p[G]}(\mathrm{Cl}(K)_p^\vee)$.
(iii) For any odd prime p write $\mathrm{Cl}(K)_p^{\vee,-}$ for the submodule of $\mathrm{Cl}(K) \otimes \mathbb{Z}_p$ upon which complex conjugation acts as multiplication by -1. Then, under a certain technical hypothesis on $\mu(K)$, the main result of Greither in [15] shows that $\mathrm{LTC}(K/k)$ also implies an explicit description of the Fitting ideal $\mathrm{Fitt}^0_{\mathbb{Z}_p[G]}(\mathrm{Cl}(K)_p^{\vee,-})$ in terms of suitably normalized Stickelberger elements. By replacing the role of 'Tate sequences for small S' in the argument of Greither by the 'T-modified Weil-étale cohomology' complexes that we introduce in §2.2 one can in fact prove the same sort of result without any hypothesis on $\mu(K)$.

1.6. New verifications of the leading term conjecture. In a sequel [9] to this article we investigate the natural Iwasawa-theoretic aspects of our general approach.

In particular, we show in [9, Th. 5.2] that, without any restriction to CM extensions (or to the 'minus parts' of conjectures), under the assumed validity of a natural main conjecture of higher rank Iwasawa theory, the validity of the p-part of $\mathrm{MRS}(L/K/k,S,T)$ for all finite abelian extensions L/k implies the validity of the p-part of $\mathrm{LTC}(K/k)$. Such a result provides an important partial converse to Theorem 1.1 and can also be used to derive new evidence in support of $\mathrm{LTC}(K/k)$.

For example, in [9, Th. 4.9] we show that, in all relevant cases, the validity of $\mathrm{MRS}(K/L/k,S,T)$ is implied by a well-known leading term formula for p-adic L-series that has been conjectured by Gross (the 'Gross-Stark conjecture' [20]). By combining this observation with significant recent work of Darmon, Dasgupta and Pollack and of Ventullo concerning the Gross-Stark conjecture we are then able to give (in [9, Cor. 5.8]) the following new evidence in support of the conjectures $\mathrm{LTC}(K/k)$ and $\mathrm{MRS}(K/L/k,S,T)$.

Theorem 1.16. *Assume that k is a totally real field, that K is an abelian CM extension of k (with maximal totally real subfield K^+) and that p is an odd prime. If the p-adic Iwasawa μ-invariant of K vanishes and at most one p-adic place of k splits in K/K^+, then for any finite subextension K'/K of the* *cyclotomic $\mathbb{Z}_p$-extension of K the minus parts of the p-parts of both* $\mathrm{LTC}(K'/k)$ *and* $\mathrm{MRS}(K'/K/k,S,T)$ *are valid.*

For examples of explicit families of extensions K/k that satisfy all of the hypotheses of Theorem 1.16 with respect to any given odd prime p see [9, Examples 5.9].

1.7. Notation. In this final subsection of the Introduction we collect together some important notation which will be used in the article.

For an abelian group G, a $\mathbb{Z}[G]$-module is simply called a G-module. Tensor products, Hom, exterior powers, and determinant modules over $\mathbb{Z}[G]$ are denoted by $\otimes_G$, Hom_G, $\bigwedge_G$, and $\det_G$, respectively. We use similar notation for Ext-groups, Fitting ideals, etc. The augmentation ideal of $\mathbb{Z}[G]$ is denoted by $I(G)$. For any G-module M and any subgroup $H \subset G$, we denote M^H for the submodule of M comprising elements fixed by H. The norm element of H is denoted by N_H, namely,

$$\mathrm{N}_H = \sum_{\sigma \in H} \sigma \in \mathbb{Z}[G].$$

Let E denote either $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$. For an abelian group A, we denote $E \otimes_{\mathbb{Z}} A$ by EA. The maximal $\mathbb{Z}$-torsion subgroup of A is denoted by A_{tors}. We write A/A_{tors} by A_{tf}. The Pontryagin dual $\mathrm{Hom}_{\mathbb{Z}}(A,\mathbb{Q}/\mathbb{Z})$ of A is denoted by $A^\vee$ for discrete A.

Fix an algebraic closure $\overline{\mathbb{Q}}$ of $\mathbb{Q}$. For a positive integer n, we denote by μ_n the group of n-th roots of unity in $\overline{\mathbb{Q}}^\times$.

Let k be a global field. The set of all infinite places of k is denoted by $S_\infty(k)$ or simply by S_∞ when k is clear from the context. (If k is a function field, then $S_\infty(k)$ is empty.) Consider a finite Galois extension K/k, and denote its Galois group by G. The set of all places of k which ramify in K is denoted by $S_{\mathrm{ram}}(K/k)$ or simply by S_{ram} when K/k is clear from the context. For any non-empty finite set S of places of k, we denote by S_K the set of places of K lying above places in S. The ring of S-integers of K is defined by

$$\mathcal{O}_{K,S} := \{a \in K : \mathrm{ord}_w(a) \geq 0 \text{ for all finite places } w \text{ of } K \text{ not contained in } S_K\},$$

where ord_w denotes the normalized additive valuation at w. The unit group of $\mathcal{O}_{K,S}$ is called the S-unit group of K. Let T be a finite set of finite places of k, which is disjoint from S. The (S,T)-unit group of K is defined by

$$\mathcal{O}_{K,S,T}^\times := \{a \in \mathcal{O}_{K,S}^\times : a \equiv 1 \pmod{w} \text{ for all } w \in T_K\}.$$

The ideal class group of $\mathcal{O}_{K,S}$ is denoted by $\mathrm{Cl}_S(K)$. This is called the S-class group of K. The (S,T)-class group of K, which we denote by $\mathrm{Cl}_S^T(K)$, is defined to be the ray class group of $\mathcal{O}_{K,S}$ modulo $\prod_{w \in T_K} w$ (namely, the quotient of the group of fractional ideals whose supports are coprime to all places above $S \cup T$ by the subgroup of principal ideals with a generator congruent to 1 modulo all places in T_K). When $S \subset S_\infty$, we omit S and write $\mathrm{Cl}^T(K)$ for $\mathrm{Cl}_S^T(K)$. When

$S \subset S_\infty$ and $T = \emptyset$, we write $\mathrm{Cl}(K)$ which is the class group of the integer ring $\mathcal{O}_K$.
We denote by $X_{K,S}$ the augmentation kernel of the divisor group $Y_{K,S} := \bigoplus_{w \in S_K} \mathbb{Z}w$. If S contains $S_\infty(k)$, then the Dirichlet regulator map

$$\lambda_{K,S} : \mathbb{R}\mathcal{O}_{K,S}^\times \longrightarrow \mathbb{R}X_{K,S},$$

defined by $\lambda_{K,S}(a) := -\sum_{w \in S_K} \log |a|_w w$, is an isomorphism.
For a place w of K, the decomposition subgroup of w in G is denoted by G_w. If w is finite, the residue field of w is denoted by $\kappa(w)$. The cardinality of $\kappa(w)$ is denoted by $\mathrm{N}w$. If the place v of k lying under w is unramified in K, then the Frobenius automorphism at w is denoted by $\mathrm{Fr}_w \in G_w$. When G is abelian, then G_w and Fr_w depend only on v, so in this case we often denote them by G_v and Fr_v respectively. The $\mathbb{C}$-linear involution $\mathbb{C}[G] \to \mathbb{C}[G]$ induced by $\sigma \mapsto \sigma^{-1}$ with $\sigma \in G$ is denoted by $x \mapsto x^\#$.
A complex of G-modules is said to be 'perfect' if it is quasi-isomorphic to a bounded complex of finitely generated projective G-modules.
We denote by $D(\mathbb{Z}[G])$ the derived category of G-modules, and by $D^{\mathrm{p}}(\mathbb{Z}[G])$ the full subcategory of $D(\mathbb{Z}[G])$ consisting of perfect complexes.

2. Canonical Selmer groups and complexes for $\mathbb{G}_m$

In this section, we give a definition of 'integral dual Selmer groups for $\mathbb{G}_m$', as analogues of Mazur-Tate's 'integral Selmer groups' defined for abelian varieties in [38]. We shall also review the construction of certain natural arithmetic complexes, which are used for the formulation of the leading term conjecture.

2.1. Integral dual Selmer groups.

Let K/k be a finite Galois extension of global fields with Galois group G. Let S be a non-empty finite set of places which contains $S_\infty(k)$. Let T be a finite set of places of k which is disjoint from S.

Definition 2.1. We define the '(S-relative T-trivialized) integral dual Selmer group for $\mathbb{G}_m$' by setting

$$\mathcal{S}_{S,T}(\mathbb{G}_{m/K}) := \mathrm{coker}(\prod_{w \notin S_K \cup T_K} \mathbb{Z} \longrightarrow \mathrm{Hom}_\mathbb{Z}(K_T^\times, \mathbb{Z})),$$

where $K_T^\times$ is the subgroup of $K^\times$ defined by

$$K_T^\times := \{a \in K^\times : \mathrm{ord}_w(a-1) > 0 \text{ for all } w \in T_K\},$$

and the homomorphism on the right hand side is defined by

$$(x_w)_w \mapsto (a \mapsto \sum_{w \notin S_K \cup T_K} \mathrm{ord}_w(a) x_w).$$

When T is empty, we omit the subscript T from this notation.

By the following proposition we see that our integral dual Selmer groups are actually like usual dual Selmer groups (see also Remark 2.3 below).

Proposition 2.2. *There is a canonical exact sequence*

$$0 \longrightarrow \mathrm{Cl}_S^T(K)^\vee \longrightarrow \mathcal{S}_{S,T}(\mathbb{G}_{m/K}) \longrightarrow \mathrm{Hom}_{\mathbb{Z}}(\mathcal{O}_{K,S,T}^\times, \mathbb{Z}) \longrightarrow 0$$

of the form (1) in §1.

Proof. Consider the commutative diagram

$$\begin{tikzcd}
0 \arrow[r] & \prod_{w\notin S_K\cup T_K}\mathbb{Z} \arrow[r] \arrow[d] & \prod_{w\notin S_K\cup T_K}\mathbb{Q} \arrow[r] \arrow[d] & \prod_{w\notin S_K\cup T_K}\mathbb{Q}/\mathbb{Z} \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & \mathrm{Hom}_{\mathbb{Z}}(K_T^\times, \mathbb{Z}) \arrow[r] & \mathrm{Hom}_{\mathbb{Z}}(K_T^\times, \mathbb{Q}) \arrow[r] & (K_T^\times)^\vee, &
\end{tikzcd}$$

where each row is the natural exact sequence, and each vertical arrow is given by $(x_w)_w \mapsto (a \mapsto \sum_{w\notin S_K\cup T_K}\mathrm{ord}_w(a)x_w)$. Using the exact sequence

$$0 \longrightarrow \mathcal{O}_{K,S,T}^\times \longrightarrow K_T^\times \stackrel{\bigoplus \mathrm{ord}_w}{\longrightarrow} \bigoplus_{w\notin S_K\cup T_K}\mathbb{Z} \longrightarrow \mathrm{Cl}_S^T(K) \longrightarrow 0$$

and applying the snake lemma to the above commutative diagram, we obtain the exact sequence

$$0 \longrightarrow \mathrm{Cl}_S^T(K)^\vee \longrightarrow \mathcal{S}_{S,T}(\mathbb{G}_{m/K}) \longrightarrow \mathrm{Hom}_{\mathbb{Z}}(\mathcal{O}_{K,S,T}^\times, \mathbb{Q}) \longrightarrow (\mathcal{O}_{K,S,T}^\times)^\vee.$$

Since the kernel of the last map is $\mathrm{Hom}_{\mathbb{Z}}(\mathcal{O}_{K,S,T}^\times, \mathbb{Z})$, we obtain the desired conclusion. □

Remark 2.3.
(i) The Bloch-Kato Selmer group $H_f^1(K, \mathbb{Q}/\mathbb{Z}(1))$ is defined to be the kernel of the diagonal map

$$K^\times \otimes \mathbb{Q}/\mathbb{Z} \longrightarrow \bigoplus_w K_w^\times/\mathcal{O}_{K_w}^\times \otimes \mathbb{Q}/\mathbb{Z}$$

where w runs over all finite places, and so lies in a canonical exact sequence

$$0 \longrightarrow \mathcal{O}_K^\times \otimes \mathbb{Q}/\mathbb{Z} \longrightarrow H_f^1(K, \mathbb{Q}/\mathbb{Z}(1)) \longrightarrow \mathrm{Cl}(K) \longrightarrow 0.$$

Its Pontryagin dual $H_f^1(K, \mathbb{Q}/\mathbb{Z}(1))^\vee$ is a finitely generated $\hat{\mathbb{Z}}$-module and our integral dual Selmer group $\mathcal{S}_{S_\infty}(\mathbb{G}_{m/K})$ provides a canonical finitely generated $\mathbb{Z}$-structure on $H_f^1(K, \mathbb{Q}/\mathbb{Z}(1))^\vee$.
(ii) In general, the exact sequence (1) also means that $\mathcal{S}_{S,T}(\mathbb{G}_{m/K})$ is a natural analogue (relative to S and T) for $\mathbb{G}_m$ over K of the 'integral Selmer group' that is defined for abelian varieties by Mazur and Tate in [38, p.720].

In the next subsection we shall give a natural cohomological interpretation of the group $\mathcal{S}_{S,T}(\mathbb{G}_{m/K})$ (see Proposition 2.4(iii)) and also show that it has a canonical 'transpose' (see Definition 2.6).

2.2. 'Weil-étale cohomology' complexes. In the following, we construct two canonical complexes of G-modules, and use them to show that there exists a canonical transpose of the module $\mathcal{S}_{S,T}(\mathbb{G}_{m/K})$. The motivation for our choice of notation (and terminology) for these complexes is explained in Remark 2.5 below.

We fix data $K/k, G, S, T$ as in the previous subsection. We also write $\mathbb{F}_{T_K}^\times$ for the direct sum $\bigoplus_{w \in T_K} \kappa(w)^\times$ of the multiplicative groups of the residue fields of all places in T_K.

Proposition 2.4. *There exist canonical complexes of G-modules $R\Gamma_c((\mathcal{O}_{K,S})_{\mathcal{W}}, \mathbb{Z})$ and $R\Gamma_{c,T}((\mathcal{O}_{K,S})_{\mathcal{W}}, \mathbb{Z})$ which satisfy all of the following conditions.*

(i) *There exists a canonical commutative diagram of exact triangles in $D(\mathbb{Z}[G])$*

(6)
$$\begin{tikzcd}
\mathrm{Hom}_{\mathbb{Z}}(\mathcal{O}_{K,S}^\times, \mathbb{Q})[-3] \arrow[r, "\theta"] \arrow[d] & R\Gamma_c(\mathcal{O}_{K,S}, \mathbb{Z}) \arrow[r] \arrow[d, equal] & R\Gamma_c((\mathcal{O}_{K,S})_{\mathcal{W}}, \mathbb{Z}) \arrow[r] \arrow[d] & {} \\
(\mathrm{Hom}_{\mathbb{Z}}(\mathcal{O}_{K,S}^\times, \mathbb{Q}) \oplus (\mathbb{F}_{T_K}^\times)^\vee)[-3] \arrow[r, "\theta'"] \arrow[d] & R\Gamma_c(\mathcal{O}_{K,S}, \mathbb{Z}) \arrow[r] & R\Gamma_{c,T}((\mathcal{O}_{K,S})_{\mathcal{W}}, \mathbb{Z}) \arrow[r] \arrow[d] & {} \\
(\mathbb{F}_{T_K}^\times)^\vee[-3] \arrow[d] & & (\mathbb{F}_{T_K}^\times)^\vee[-2] \arrow[d, "\theta''"] & \\
{} & & {} &
\end{tikzcd}$$

in which the first column is induced by the obvious exact sequence

$$0 \longrightarrow \mathrm{Hom}_{\mathbb{Z}}(\mathcal{O}_{K,S}^\times, \mathbb{Q}) \longrightarrow \mathrm{Hom}_{\mathbb{Z}}(\mathcal{O}_{K,S}^\times, \mathbb{Q}) \oplus (\mathbb{F}_{T_K}^\times)^\vee \longrightarrow (\mathbb{F}_{T_K}^\times)^\vee \longrightarrow 0$$

and $H^2(\theta'')$ is the Pontryagin dual of the natural injective homomorphism

$$H^3(R\Gamma_c((\mathcal{O}_{K,S})_{\mathcal{W}}, \mathbb{Z}))^\vee = \mathcal{O}_{K,\mathrm{tors}}^\times \longrightarrow \mathbb{F}_{T_K}^\times.$$

(ii) *If S' is a set of places of k which contains S and is disjoint from T, then there is a canonical exact triangle in $D(\mathbb{Z}[G])$*

$$R\Gamma_{c,T}((\mathcal{O}_{K,S'})_{\mathcal{W}}, \mathbb{Z}) \longrightarrow R\Gamma_{c,T}((\mathcal{O}_{K,S})_{\mathcal{W}}, \mathbb{Z}) \longrightarrow \bigoplus_{w \in S'_K \setminus S_K} R\Gamma((\kappa(w))_{\mathcal{W}}, \mathbb{Z}),$$

where $R\Gamma((\kappa(w))_{\mathcal{W}}, \mathbb{Z})$ is the complex of G_w-modules which lies in the exact triangle

$$\mathbb{Q}[-2] \longrightarrow R\Gamma(\kappa(w), \mathbb{Z}) \longrightarrow R\Gamma((\kappa(w))_{\mathcal{W}}, \mathbb{Z}) \longrightarrow,$$

where the H^2 of the first arrow is the natural map

$$\mathbb{Q} \longrightarrow \mathbb{Q}/\mathbb{Z} = H^2(\kappa(w), \mathbb{Z}).$$

(iii) *The complex* $R\Gamma_{c,T}((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{Z})$ *is acyclic outside degrees one, two and three, and there are canonical isomorphisms*

$$H^i(R\Gamma_{c,T}((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{Z}))\simeq\begin{cases}Y_{K,S}/\Delta_S(\mathbb{Z}) & \text{if } i=1,\\ \mathcal{S}_{S,T}(\mathbb{G}_{m/K}) & \text{if } i=2,\\ (K^\times_{T,\mathrm{tors}})^\vee & \text{if } i=3,\end{cases}$$

where Δ_S *is the natural diagonal map.*

(iv) *If* S *contains* $S_{\mathrm{ram}}(K/k)$, *then* $R\Gamma_c((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{Z})$ *and* $R\Gamma_{c,T}((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{Z})$ *are both perfect complexes of* G*-modules.*

Proof. In this argument we use the fact that morphisms in $D(\mathbb{Z}[G])$ between bounded above complexes $K_1^\bullet$ and $K_2^\bullet$ can be computed by means of the spectral sequence

$$E_2^{p,q}=\prod_{a\in\mathbb{Z}}\mathrm{Ext}^p_G(H^a(K_1^\bullet),H^{q+a}(K_2^\bullet))\Rightarrow H^{p+q}(R\,\mathrm{Hom}_G(K_1^\bullet,K_2^\bullet)) \tag{7}$$

constructed by Verdier in [53, III, 4.6.10].

Set $C^\bullet=C^\bullet_S:=R\Gamma_c(\mathcal{O}_{K,S},\mathbb{Z})$ and $W:=\mathrm{Hom}_{\mathbb{Z}}(\mathcal{O}^\times_{K,S},\mathbb{Q})$ for simplicity. Then we recall first that $C^\bullet$ is acyclic outside degrees one, two and three, that there are canonical isomorphisms

$$H^i(C^\bullet)\cong\begin{cases}Y_{K,S}/\Delta_S(\mathbb{Z}) & \text{if } i=1,\\ \mathrm{Cl}_S(K)^\vee & \text{if } i=2,\\ (\mathcal{O}^\times_{K,S})^\vee & \text{if } i=3,\end{cases} \tag{8}$$

where Δ_S is the map that occurs in the statement of claim (iii) and that, when S contains $S_{\mathrm{ram}}(K/k)$, $C^\bullet$ is isomorphic to a bounded complex of cohomologically-trivial G-modules.

It is not difficult to see that the groups $\mathrm{Ext}^i_G(W,H^{3-i}(C^\bullet))$ vanish for all $i>0$, and so the spectral sequence (7) implies that the 'passage to cohomology' homomorphism

$$H^0(R\,\mathrm{Hom}_G(W[-3],C^\bullet))=\mathrm{Hom}_{D(\mathbb{Z}[G])}(W[-3],C^\bullet)\longrightarrow\mathrm{Hom}_G(W,(\mathcal{O}^\times_{K,S})^\vee)$$

is bijective. We may therefore define θ to be the unique morphism in $D(\mathbb{Z}[G])$ for which $H^3(\theta)$ is equal to the natural map

$$W=\mathrm{Hom}_{\mathbb{Z}}(\mathcal{O}^\times_{K,S},\mathbb{Q})\longrightarrow\mathrm{Hom}_{\mathbb{Z}}(\mathcal{O}^\times_{K,S},\mathbb{Q}/\mathbb{Z})=(\mathcal{O}^\times_{K,S})^\vee$$

and then take $C^\bullet_{\mathcal{W}}:=R\Gamma_c((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{Z})$ to be any complex which lies in an exact triangle of the form that occurs in the upper row of (6). An analysis of the long exact cohomology sequence of this triangle then shows that $C^\bullet_{\mathcal{W}}$ is acyclic outside degrees one, two and three, that $H^1(C^\bullet_{\mathcal{W}})=H^1(C^\bullet)$, that $H^2(C^\bullet_{\mathcal{W}})_{\mathrm{tors}}=H^2(C^\bullet)$, that $H^2(C^\bullet_{\mathcal{W}})_{\mathrm{tf}}=\mathrm{Hom}_{\mathbb{Z}}(\mathcal{O}^\times_{K,S},\mathbb{Z})$ and that $H^3(C^\bullet_{\mathcal{W}})=(\mathcal{O}^\times_{K,\mathrm{tors}})^\vee$. In particular, when S contains $S_{\mathrm{ram}}(K/k)$, since each of these groups is finitely generated and both of the complexes $W[-3]$ and $C^\bullet$ are represented by bounded complexes of cohomologically-trivial G-modules, this implies that $C^\bullet_{\mathcal{W}}$ is perfect.

To define the morphism θ' we first choose a finite set S'' of places of k which is disjoint from $S \cup T$ and such that $\mathrm{Cl}_{S'}(K)$ vanishes for $S' := S \cup S''$. Note that (8) with S replaced by S' implies $C^\bullet_{S'}$ is acyclic outside degrees one and three. We also note that, since each place in T is unramified in K/k, there is also an exact sequence of G-modules

$$0 \longrightarrow \bigoplus_{v\in T} \mathbb{Z}[G] \xrightarrow{(1-\mathrm{N}v\mathrm{Fr}_w)_v} \bigoplus_{v\in T} \mathbb{Z}[G] \longrightarrow (\mathbb{F}^\times_{T_K})^\vee \longrightarrow 0 \tag{9}$$

where w is any choice of place of K above v. This sequence shows both that $(\mathbb{F}^\times_{T_K})^\vee[-3]$ is a perfect complex of G-modules and also that the functor $\mathrm{Ext}^i_G((\mathbb{F}^\times_{T_K})^\vee, -)$ vanishes for all $i > 1$. In particular, the spectral sequence (7) implies that in this case the passage to cohomology homomorphism

$$\mathrm{Hom}_{D(\mathbb{Z}[G])}((\mathbb{F}^\times_{T_K})^\vee[-3], C^\bullet_{S'}) \longrightarrow \mathrm{Hom}_G((\mathbb{F}^\times_{T_K})^\vee, (\mathcal{O}^\times_{K,S'})^\vee)$$

is bijective. We may therefore define θ' to be the morphism which restricts on $W[-3]$ to give θ and on $(\mathbb{F}^\times_{T_K})^\vee[-3]$ to give the composite morphism

$$(\mathbb{F}^\times_{T_K})^\vee[-3] \xrightarrow{\theta'_1} R\Gamma_c(\mathcal{O}_{K,S'}, \mathbb{Z}) \xrightarrow{\theta'_2} R\Gamma_c(\mathcal{O}_{K,S}, \mathbb{Z})$$

where θ'_1 is the unique morphism for which $H^3(\theta'_1)$ is the Pontryagin dual of the natural map $\mathcal{O}^\times_{K,S'} \to \mathbb{F}^\times_{T_K}$ and θ'_2 occurs in the canonical exact triangle

$$R\Gamma_c(\mathcal{O}_{K,S'}, \mathbb{Z}) \xrightarrow{\theta'_2} R\Gamma_c(\mathcal{O}_{K,S}, \mathbb{Z}) \longrightarrow \bigoplus_{w\in S''_K} R\Gamma(\kappa(w), \mathbb{Z}) \longrightarrow \tag{10}$$

constructed by Milne in [39, Chap. II, Prop. 2.3 (d)].

We now take $C^\bullet_{\mathcal{W},T} := R\Gamma_{c,T}((\mathcal{O}_{K,S})_\mathcal{W}, \mathbb{Z})$ to be any complex which lies in an exact triangle of the form that occurs in the second row of (6) and then, just as above, an analysis of this triangle shows that $C^\bullet_{\mathcal{W},T}$ is a perfect complex of G-modules when S contains $S_{\mathrm{ram}}(K/k)$. Note also that since for this choice of θ' the upper left hand square of (6) commutes the diagram can then be completed to give the right hand vertical exact triangle. The claim (ii) follows easily from the above constructions.

It only remains to prove claim (iii). It is easy to see that the groups $H^i(R\Gamma_{c,T}((\mathcal{O}_{K,S})_\mathcal{W}, \mathbb{Z}))$ for $i = 1$ and 3 are as described in claim (iii), so we need only prove that there is a natural isomorphism

$$H^2(R\Gamma_{c,T}((\mathcal{O}_{K,S})_\mathcal{W}, \mathbb{Z})) \simeq \mathcal{S}_{S,T}(\mathbb{G}_{m/K}).$$

To do this we first apply claim (ii) for a set S' that is large enough to ensure that $\mathrm{Cl}^T_{S'}(K)$ vanishes. Since in this case

$$H^2(R\Gamma_{c,T}((\mathcal{O}_{K,S'})_\mathcal{W}, \mathbb{Z})) = \mathrm{Hom}_\mathbb{Z}(\mathcal{O}^\times_{K,S',T}, \mathbb{Z}),$$

we obtain in this way a canonical isomorphism

$$H^2(R\Gamma_{c,T}((\mathcal{O}_{K,S})_\mathcal{W}, \mathbb{Z})) \simeq \mathrm{coker}(\bigoplus_{w\in S'_K\setminus S_K} \mathbb{Z} \longrightarrow \mathrm{Hom}_\mathbb{Z}(\mathcal{O}^\times_{K,S',T}, \mathbb{Z})). \tag{11}$$

Consider next the commutative diagram

$$\begin{tikzcd}
0 \arrow[r] & \prod_{w\notin S'_K\cup T_K}\mathbb{Z} \arrow[r] \arrow[d, equal] & \prod_{w\notin S_K\cup T_K}\mathbb{Z} \arrow[r] \arrow[d] & \bigoplus_{w\in S'_K\setminus S_K}\mathbb{Z} \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & \prod_{w\notin S'_K\cup T_K}\mathbb{Z} \arrow[r] & \mathrm{Hom}_{\mathbb{Z}}(K_T^\times,\mathbb{Z}) \arrow[r] & \mathrm{Hom}_{\mathbb{Z}}(\mathcal{O}_{K,S',T}^\times,\mathbb{Z}) \arrow[r] & 0
\end{tikzcd}$$

with exact rows, where the first exact row is the obvious one, the second is the dual of the exact sequence

$$0 \longrightarrow \mathcal{O}_{K,S',T}^\times \longrightarrow K_T^\times \stackrel{\bigoplus \mathrm{ord}_w}{\longrightarrow} \bigoplus_{w\notin S'_K\cup T_K}\mathbb{Z} \longrightarrow 0,$$

and the vertical arrows are given by $(x_w)_w \mapsto (a \mapsto \sum_w \mathrm{ord}_w(a)x_w)$. From this we have the canonical isomorphism

$$\mathcal{S}_{S,T}(\mathbb{G}_{m/K}) \simeq \mathrm{coker}(\bigoplus_{w\in S'_K\setminus S_K}\mathbb{Z} \longrightarrow \mathrm{Hom}_{\mathbb{Z}}(\mathcal{O}_{K,S',T}^\times,\mathbb{Z})). \tag{12}$$

From (11) and (12) our claim follows. □

Given the constructions in Proposition 2.4, in each degree i we set

$$H^i_{c,T}((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{Z}) := H^i(R\Gamma_{c,T}((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{Z})).$$

We also define a complex

$$R\Gamma_T((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{G}_m) := R\,\mathrm{Hom}_{\mathbb{Z}}(R\Gamma_{c,T}((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{Z}),\mathbb{Z})[-2].$$

We endow this complex with the natural contragredient action of G and then in each degree i set

$$H^i_T((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{G}_m) := H^i(R\Gamma_T((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{G}_m)).$$

Remark 2.5. Our notation for the above cohomology groups and complexes is motivated by the following facts.
(i) Assume that k is a function field. Write C_k for the corresponding smooth projective curve, $C_{k,W\text{ét}}$ for the Weil-étale site on C_k that is defined by Lichtenbaum in [33, §2] and j for the open immersion $\mathrm{Spec}(\mathcal{O}_{k,S}) \longrightarrow C_k$. Then the group $H^i_c((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{Z})$ defined above is canonically isomorphic to the Weil-étale cohomology group $H^i(C_{k,W\text{ét}}, j_!\mathbb{Z})$.
(ii) Assume that k is a number field. In this case there has as yet been no construction of a 'Weil-étale topology' for $Y_S := \mathrm{Spec}(\mathcal{O}_{K,S})$ with all of the properties that are conjectured by Lichtenbaum in [34]. However, if $\overline{Y_S}$ is a compactification of Y_S and ϕ is the natural inclusion $Y_S \subset \overline{Y_S}$, then the approach of [4] can be used to show that, should such a topology exist with all of the expected properties, then the groups $H^i_c((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{Z})$ defined above would be canonically isomorphic to the group $H^i_c(Y_S,\mathbb{Z}) := H^i(\overline{Y_S},\phi_!\mathbb{Z})$ that is discussed in [34].
(iii) The definition of $R\Gamma_T((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{G}_m)$ as the (shifted) linear dual of the complex $R\Gamma_{c,T}((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{Z})$ is motivated by [4, Rem. 3.8] and hence by the

duality theorem in Weil-étale cohomology for curves over finite fields that is proved by Lichtenbaum in [33].

An analysis of the complex $R\Gamma_{c,T}((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{Z})$ as in the proof of Lemma 2.8 below then leads us to give the following definition. In this definition we use the notion of 'transpose' in the sense of Jannsen's homotopy theory of modules [23].

Definition 2.6. The 'transpose' of $\mathcal{S}_{S,T}(\mathbb{G}_{m/K})$ is the group

$$\mathcal{S}^{\mathrm{tr}}_{S,T}(\mathbb{G}_{m/K}) := H^1_T((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{G}_m) = H^{-1}(R\operatorname{Hom}_{\mathbb{Z}}(R\Gamma_{c,T}((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{Z}),\mathbb{Z})).$$

When T is empty, we omit the subscript T from this notation.

Remark 2.7. By using the spectral sequence

$$E_2^{p,q} = \operatorname{Ext}^p_{\mathbb{Z}}(H^{-q}_{c,T}((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{Z}),\mathbb{Z}) \Rightarrow H^{p+q+2}_T((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{G}_m),$$

which is obtained from (7), one can check that $R\Gamma_T((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{G}_m)$ is acyclic outside degrees zero and one, that there is a canonical isomorphism

$$H^0_T((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{G}_m) \simeq \mathcal{O}^\times_{K,S,T},$$

and that there is a canonical exact sequence

$$0 \longrightarrow \mathrm{Cl}^T_S(K) \longrightarrow \mathcal{S}^{\mathrm{tr}}_{S,T}(\mathbb{G}_{m/K}) \longrightarrow X_{K,S} \longrightarrow 0$$

of the form (2) in §1.

In the sequel we shall say that a G-module M has a 'locally-quadratic presentation' if it lies in an exact sequence of finitely generated G-modules of the form

$$P \to P' \to M \to 0$$

in which P and P' are projective and the $\mathbb{Q}[G]$-modules $\mathbb{Q}P$ and $\mathbb{Q}P'$ are isomorphic.

Lemma 2.8. *Assume that G is abelian, that S contains $S_\infty(k) \cup S_{\mathrm{ram}}(K/k)$, and that $\mathcal{O}^\times_{K,S,T}$ is $\mathbb{Z}$-torsion-free. Then each of the groups $\mathcal{S}_{S,T}(\mathbb{G}_{m/K})$ and $\mathcal{S}^{\mathrm{tr}}_{S,T}(\mathbb{G}_{m/K})$ have locally-quadratic presentations, and for each non-negative integer i one has an equality*

$$\mathrm{Fitt}^i_G(\mathcal{S}^{\mathrm{tr}}_{S,T}(\mathbb{G}_{m/K})) = \mathrm{Fitt}^i_G(\mathcal{S}_{S,T}(\mathbb{G}_{m/K}))^{\#}.$$

Proof. Set

$$C^\bullet := R\Gamma_{c,T}((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{Z})$$

and

$$C^{\bullet,*} := R\operatorname{Hom}_{\mathbb{Z}}(R\Gamma_{c,T}((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{Z}),\mathbb{Z}).$$

From Proposition 2.4 we also know that $C^\bullet$ is a perfect complex of G-modules that is acyclic outside degree one and two and $\mathbb{Z}$-torsion-free in degree one. This implies, by a standard argument, that $C^\bullet$ can be represented by a complex $P \xrightarrow{\delta} P'$ of G-modules, where P and P' are finitely generated and projective and the first term is placed in degree one, and hence that there is a tautological exact sequence of G-modules

$$0 \longrightarrow H^1(C^\bullet) \longrightarrow P \xrightarrow{\delta} P' \longrightarrow H^2(C^\bullet) \longrightarrow 0. \tag{13}$$

The descriptions in Proposition 2.4(iii) imply that the linear dual of the Dirichlet regulator map $\lambda_{K,S}$ induces an isomorphism of $\mathbb{R}[G]$-modules

$$\lambda^*_{K,S} : \mathbb{R}H^1(C^\bullet) \cong \mathbb{R}H^2(C^\bullet). \tag{14}$$

Taken in conjunction with the sequence (13) this isomorphism implies that the $\mathbb{Q}[G]$-modules $\mathbb{Q}P$ and $\mathbb{Q}P'$ are isomorphic and hence that $\mathcal{S}_{S,T}(\mathbb{G}_{m/K}) = H^2(C^\bullet)$ has a locally-quadratic presentation, as claimed.

The complex $C^{\bullet,*}[-2]$ is represented by $\mathrm{Hom}_\mathbb{Z}(P',\mathbb{Z}) \xrightarrow{\delta^*} \mathrm{Hom}_\mathbb{Z}(P,\mathbb{Z})$ where the linear duals are endowed with contragredient action of G, the first term is placed in degree zero and δ^* is the map induced by δ. There is therefore a tautological exact sequence

$$0 \longrightarrow H^0(C^{\bullet,*}[-2]) \longrightarrow \mathrm{Hom}_\mathbb{Z}(P',\mathbb{Z}) \xrightarrow{\delta^*} \mathrm{Hom}_\mathbb{Z}(P,\mathbb{Z}) \longrightarrow H^1(C^{\bullet,*}[-2]) \longrightarrow 0, \tag{15}$$

and, since the above observations imply that $\mathrm{Hom}_\mathbb{Z}(P',\mathbb{Z})$ and $\mathrm{Hom}_\mathbb{Z}(P,\mathbb{Z})$ are projective G-modules that span isomorphic $\mathbb{Q}[G]$-spaces, this sequence implies that the module $\mathcal{S}^{\mathrm{tr}}_{S,T}(\mathbb{G}_{m/K}) = H^1(C^{\bullet,*}[-2])$ has a locally-quadratic presentation.

It now only remains to prove the final claim and it is enough to prove this after completion at each prime p. We shall denote for any abelian group A the p-completion $A\otimes\mathbb{Z}_p$ of A by A_p. By Swan's Theorem (cf. [11, (32.1)]) one knows that for each prime p the $\mathbb{Z}_p[G]$-modules P_p and P'_p are both free of rank, d say, that is independent of p. In particular, after fixing bases of P_p and P'_p the homomorphism $P_p \xrightarrow{\delta} P'_p$ corresponds to a matrix $A_{\delta,p}$ in $\mathrm{M}_d(\mathbb{Z}_p[G])$ and the sequence (13) implies that the ideal $\mathrm{Fitt}^i_G(H^2(C^\bullet))_p$ is generated over $\mathbb{Z}_p[G]$ by the determinants of all $(d-i)\times(d-i)$ minors of $A_{\delta,p}$. The corresponding dual bases induce identifications of both $\mathrm{Hom}_\mathbb{Z}(P',\mathbb{Z})_p$ and $\mathrm{Hom}_\mathbb{Z}(P,\mathbb{Z})_p$ with $\mathbb{Z}_p[G]^{\oplus d}$, with respect to which the homomorphism $\mathrm{Hom}_\mathbb{Z}(P',\mathbb{Z})_p \xrightarrow{\delta^*} \mathrm{Hom}_\mathbb{Z}(P,\mathbb{Z})_p$ is represented by the matrix $A^{\mathrm{tr},\#}_{\delta,p}$ that is obtained by applying the involution # to each entry of the transpose of $A_{\delta,p}$. The exact sequence (15) therefore implies that $\mathrm{Fitt}^i_G(H^1(C^{\bullet,*}[-2]))_p$ is generated over $\mathbb{Z}_p[G]$ by the determinants of all $(d-i)\times(d-i)$ minors of $A^{\mathrm{tr},\#}_{\delta,p}$. Hence one has an equality

$$\mathrm{Fitt}^i_G(H^2(C^\bullet))_p = \mathrm{Fitt}^i_G(H^1(C^{\bullet,*}[-2]))^\#_p,$$

as required. □

2.3. Tate sequences. In this subsection we review the construction of Tate's exact sequence, which is used in the formulation of the leading term conjecture in the next section. Let $K/k, G, S$ be as in the previous subsection. We assume that $S_{\mathrm{ram}}(K/k) \subset S$. We assume only in this subsection that S is large enough so that $\mathrm{Cl}_S(K)$ vanishes.

In this setting, Tate constructed a 'fundamental class' $\tau_{K/k,S} \in \mathrm{Ext}^2_G(X_{K,S}, \mathcal{O}^\times_{K,S})$ using the class field theory [49]. This class $\tau_{K/k,S}$ has the following property: if we regard $\tau_{K/k,S}$ as an element of $H^2(G, \mathrm{Hom}_{\mathbb{Z}}(X_{K,S}, \mathcal{O}^\times_{K,S}))$ via the canonical isomorphism

$$\begin{aligned}\mathrm{Ext}^2_G(X_{K,S}, \mathcal{O}^\times_{K,S}) \simeq \mathrm{Ext}^2_G(\mathbb{Z}, \mathrm{Hom}_{\mathbb{Z}}(X_{K,S}, \mathcal{O}^\times_{K,S}))\\ = H^2(G, \mathrm{Hom}_{\mathbb{Z}}(X_{K,S}, \mathcal{O}^\times_{K,S})),\end{aligned}$$

then, for every integer i, the map between Tate cohomology groups

$$\widehat{H}^i(G, X_{K,S}) \xrightarrow{\sim} \widehat{H}^{i+2}(G, \mathcal{O}^\times_{K,S})$$

that is defined by taking cup product with $\tau_{K/k,S}$ is bijective.
The Yoneda extension class of $\tau_{K/k,S}$ is therefore represented by an exact sequence of the following sort:

$$0 \longrightarrow \mathcal{O}^\times_{K,S} \longrightarrow A \longrightarrow B \longrightarrow X_{K,S} \longrightarrow 0, \tag{16}$$

where A and B are finitely generated cohomologically-trivial G-modules (see [50, Chap. II, Th. 5.1]). We call this sequence a 'Tate sequence' for K/k.

Proposition 2.9. *The complex $R\Gamma((\mathcal{O}_{K,S})_{\mathcal{W}}, \mathbb{G}_m)$ defines an element of*

$$\mathrm{Ext}^2_G(\mathcal{S}^{\mathrm{tr}}_S(\mathbb{G}_{m/K}), \mathcal{O}^\times_{K,S}).$$

This element is equal to Tate's fundamental class $\tau_{K/k,S}$.

Proof. The first assertion follows directly from the discussion of Remark 2.7. The assumed vanishing of $\mathrm{Cl}_S(K)$ combines with the exact sequence (2) to imply that $\mathcal{S}^{\mathrm{tr}}_S(\mathbb{G}_{m/K}) = X_{K,S}$. Given this, the second claim is proved by the first author in [4, Prop. 3.5(f)] □

3. Zeta elements and the leading term conjecture

In this section, we suppose that K/k is a finite abelian extension of global fields with Galois group G.
We fix a finite non-empty set of places S of k which contains both $S_\infty(k)$ and $S_{\mathrm{ram}}(K/k)$ and an auxiliary finite set of places T of k that is disjoint from S.

3.1. L-functions. We recall the definition of (abelian) L-functions of global fields. For any linear character $\chi \in \widehat{G} := \mathrm{Hom}(G, \mathbb{C}^\times)$, we define the S-truncated T-modified L-function for K/k and χ by setting

$$L_{k,S,T}(\chi, s) := \prod_{v \in T}(1 - \chi(\mathrm{Fr}_v)\mathrm{N}v^{1-s}) \prod_{v \notin S}(1 - \chi(\mathrm{Fr}_v)\mathrm{N}v^{-s})^{-1}.$$

This is a complex function defined on $\mathrm{Re}(s) > 1$ and is well-known to have a meromorphic continuation on $\mathbb{C}$ and to be holomorphic at $s = 0$. We denote by $r_{\chi,S}$ the order of vanishing of $L_{k,S,T}(\chi, s)$ at $s = 0$ (this clearly does not depend on T). We denote the leading coefficient of the Taylor expansion of $L_{k,S,T}(\chi, s)$ at $s = 0$ by

$$L^*_{k,S,T}(\chi, 0) := \lim_{s \to 0} s^{-r_{\chi,S}} L_{k,S,T}(\chi, s).$$

We then define the S-truncated T-modified equivariant L-function for K/k by setting

$$\theta_{K/k,S,T}(s) := \sum_{\chi \in \widehat{G}} L_{k,S,T}(\chi^{-1}, s) e_\chi,$$

where $e_\chi := \frac{1}{|G|} \sum_{\sigma \in G} \chi(\sigma)\sigma^{-1}$, and we define its leading term to be

$$\theta^*_{K/k,S,T}(0) := \sum_{\chi \in \widehat{G}} L^*_{k,S,T}(\chi^{-1}, 0) e_\chi.$$

It is then easy to see that $\theta^*_{K/k,S,T}(0)$ belongs to $\mathbb{R}[G]^\times$.
When $T = \emptyset$, we simply denote $L_{k,S,\emptyset}(\chi, s)$, $\theta_{K/k,S,\emptyset}(s)$, etc., by $L_{k,S}(\chi, s)$, $\theta_{K/k,S}(s)$, etc., respectively, and refer to them as the S-truncated L-function for K/k, S-truncated equivariant L-function for K/k, etc., respectively.

3.2. The leading term lattice. In this section we recall the explicit formulation of a conjectural description of the lattice $\theta^*_{K/k,S,T}(0)\cdot\mathbb{Z}[G]$ which involves Tate sequences. In particular, up until Remark 3.3, we always assume (without further explicit comment) that S is large enough to ensure the group $\mathrm{Cl}_S(K)$ vanishes.
At the outset we also note that, as observed by Knudsen and Mumford in [27], to avoid certain technical difficulties regarding signs, determinant modules must be regarded as graded invertible modules. Nevertheless, for simplicity of notation, in the following we have preferred to omit explicit reference to the grading of any graded invertible modules. Thus, for a finitely generated projective G-module P, we have abbreviated the graded invertible G-module $(\det_G(P), \mathrm{rk}_G(P))$ to $\det_G(P)$, where $\mathrm{rk}_G(P)$ is the rank of P. Since the notation $\det_G(P)$ explicitly indicates P, which in turn determines $\mathrm{rk}_G(P)$, we feel that this abbreviation should not cause difficulties.
We shall also use the following general notation. Suppose that we have a perfect complex $C^\bullet$ of G-modules, which is concentrated in degree i and $i+1$ with some integer i, and an isomorphism $\lambda : \mathbb{R}H^i(C^\bullet) \xrightarrow{\sim} \mathbb{R}H^{i+1}(C^\bullet)$. Then we define an isomorphism

$$\vartheta_\lambda : \mathbb{R}\mathrm{det}_G(C^\bullet) \xrightarrow{\sim} \mathbb{R}[G]$$

as follows:

$$\begin{aligned}
\mathbb{R}\mathrm{det}_G(C^\bullet) &\xrightarrow{\sim} \bigotimes_{j\in\mathbb{Z}} \mathrm{det}^{(-1)^j}_{\mathbb{R}[G]}(\mathbb{R}C^j) \\
&\xrightarrow{\sim} \bigotimes_{j\in\mathbb{Z}} \mathrm{det}^{(-1)^j}_{\mathbb{R}[G]}(\mathbb{R}H^j(C^\bullet)) \\
&= \mathrm{det}^{(-1)^i}_{\mathbb{R}[G]}(\mathbb{R}H^i(C^\bullet)) \otimes_{\mathbb{R}[G]} \mathrm{det}^{(-1)^{i+1}}_{\mathbb{R}[G]}(\mathbb{R}H^{i+1}(C^\bullet)) \\
&\xrightarrow{\sim} \mathrm{det}^{(-1)^i}_{\mathbb{R}[G]}(\mathbb{R}H^{i+1}(C^\bullet)) \otimes_{\mathbb{R}[G]} \mathrm{det}^{(-1)^{i+1}}_{\mathbb{R}[G]}(\mathbb{R}H^{i+1}(C^\bullet)) \\
&\xrightarrow{\sim} \mathbb{R}[G],
\end{aligned}$$

where the fourth isomorphism is induced by $\lambda^{(-1)^i}$.
Let A and B be the G-modules which appear in the Tate sequence (16). Since we have the regulator isomorphism

$$\lambda_{K,S} : \mathbb{R}\mathcal{O}_{K,S}^{\times} \xrightarrow{\sim} \mathbb{R}X_{K,S},$$

the above construction for $C^{\bullet} = (A \to B)$, where A is placed in degree 0, gives the isomorphism

$$\vartheta_{\lambda_{K,S}} : \mathbb{R}\mathrm{det}_G(A) \otimes_{\mathbb{R}[G]} \mathbb{R}\mathrm{det}_G^{-1}(B) \xrightarrow{\sim} \mathbb{R}[G].$$

We study the following conjecture.

Conjecture 3.1. *In $\mathbb{R}[G]$ one has*

$$\vartheta_{\lambda_{K,S}}(\mathrm{det}_G(A) \otimes_G \mathrm{det}_G^{-1}(B)) = \theta^*_{K/k,S}(0) \cdot \mathbb{Z}[G].$$

Remark 3.2. This conjecture coincides with the conjecture $\mathrm{C}(K/k)$ stated in [3, §6.3]. The observations made in [3, Rem. 6.2] therefore imply that Conjecture 3.1 is equivalent in the number field case to the 'equivariant Tamagawa number conjecture' [7, Conj. 4 (iv)] for the pair $(h^0(\mathrm{Spec}\, K), \mathbb{Z}[G])$, that the validity of Conjecture 3.1 is independent of S and of the choice of Tate sequence and that its validity for the extension K/k implies its validity for all extensions F/E with $k \subseteq E \subseteq F \subseteq K$.

Remark 3.3. Conjecture 3.1 is known to be valid in each of the following cases:

(i) K is an abelian extension of $\mathbb{Q}$ (by Greither and the first author [8] and Flach [14]),

(ii) k is a global function field (by the first author [5]),

(iii) $[K : k] \leq 2$ (by Kim [26, §2.4, Rem. i)]).

In the following result we do not assume that the group $\mathrm{Cl}_S(K)$ vanishes and we interpret the validity of Conjecture 3.1 in terms of the 'Weil-étale cohomology' complexes $R\Gamma_{c,T}((\mathcal{O}_{K,S})_{\mathcal{W}}, \mathbb{Z})$ and $R\Gamma_T((\mathcal{O}_{K,S})_{\mathcal{W}}, \mathbb{G}_m)$ defined in §2.2.
We note at the outset that $R\Gamma_{c,T}((\mathcal{O}_{K,S})_{\mathcal{W}}, \mathbb{Z})$ (resp. $R\Gamma_T((\mathcal{O}_{K,S})_{\mathcal{W}}, \mathbb{G}_m)$) is represented by a complex which is concentrated in degrees one and two (resp. zero and one), and so we can define the isomorphism

$$\vartheta_{\lambda^*_{K,S}} : \mathbb{R}\mathrm{det}_G(R\Gamma_{c,T}((\mathcal{O}_{K,S})_{\mathcal{W}}, \mathbb{Z})) \xrightarrow{\sim} \mathbb{R}[G]$$

$$(\text{resp. } \vartheta_{\lambda_{K,S}} : \mathbb{R}\mathrm{det}_G(R\Gamma_T((\mathcal{O}_{K,S})_{\mathcal{W}}, \mathbb{G}_m)) \xrightarrow{\sim} \mathbb{R}[G]).$$

Proposition 3.4. *Let S be any finite non-empty set of places of k containing both $S_\infty(k)$ and $S_{\mathrm{ram}}(K/k)$ and let T be any finite set of places of k that is disjoint from S. Then the following conditions on K/k are equivalent.*

(i) *Conjecture 3.1 is valid.*

(ii) *In $\mathbb{R}[G]$ one has an equality*

$$\vartheta_{\lambda^*_{K,S}}(\mathrm{det}_G(R\Gamma_{c,T}((\mathcal{O}_{K,S})_{\mathcal{W}}, \mathbb{Z}))) = \theta^*_{K/k,S,T}(0)^{-1\#} \cdot \mathbb{Z}[G].$$

(iii) *In $\mathbb{R}[G]$ one has an equality*

$$\vartheta_{\lambda_{K,S}}(\mathrm{det}_G(R\Gamma_T((\mathcal{O}_{K,S})_{\mathcal{W}}, \mathbb{G}_m))) = \theta^*_{K/k,S,T}(0) \cdot \mathbb{Z}[G].$$

Proof. For any finitely generated projective G-module P of (constant) rank d there is a natural identification

$$\bigwedge_G^d \mathrm{Hom}_{\mathbb{Z}}(P,\mathbb{Z}) \cong \bigwedge_G^d \mathrm{Hom}_G(P,\mathbb{Z}[G])^{\#} \cong \mathrm{Hom}_G(\bigwedge_G^d P,\mathbb{Z}[G])^{\#},$$

where G acts on $\mathrm{Hom}_{\mathbb{Z}}(P,\mathbb{Z})$ contragrediently and on $\mathrm{Hom}_G(P,\mathbb{Z}[G])$ via right multiplication. The equivalence of the equalities in claims (ii) and (iii) is therefore a consequence of the fact that for any element Δ of the mutliplicative group of invertible $\mathbb{Z}[G]$-lattices in $\mathbb{R}[G]$ the evaluation pairing identifies $\mathrm{Hom}_G(\Delta,\mathbb{Z}[G])^{\#}$ with the image under the involution $\#$ of the inverse lattice Δ^{-1}.

To relate the equalities in claims (ii) and (iii) to Conjecture 3.1 we note first that the third column of (6) implies that

$$\begin{aligned}&\vartheta_{\lambda^*_{K,S}}(\mathrm{det}_G(R\Gamma_{c,T}((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{Z})))\\ &= \mathrm{det}_G((\mathbb{F}^{\times}_{T_K})^{\vee}[-2])\cdot\vartheta_{\lambda^*_{K,S}}(\mathrm{det}_G(R\Gamma_c((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{Z}))),\end{aligned}$$

whilst the resolution (9) implies that

$$\begin{aligned}\mathrm{det}_G((\mathbb{F}^{\times}_{T_K})^{\vee}[-2]) &= (\prod_{v\in T}(1-\mathrm{N}v\mathrm{Fr}_w))^{-1}\cdot\mathbb{Z}[G]\\ &= (\theta^*_{K/k,S,T}(0)/\theta^*_{K/k,S}(0))^{-1\#}\cdot\mathbb{Z}[G].\end{aligned}$$

The equality in claim (ii) is therefore equivalent to an equality

$$\vartheta_{\lambda^*_{K,S}}(\mathrm{det}_G(R\Gamma_c((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{Z}))) = \theta^*_{K/k,S}(0)^{-1\#}\cdot\mathbb{Z}[G]. \tag{17}$$

We now choose an auxiliary set of places S'' as in the proof of Proposition 2.4 and set $S' := S\cup S''$. By Chebotarev density theorem we can even assume that all places in S'' split completely in K/k and, for simplicity, this is what we shall do. Then, in this case, the exact triangle (10) combines with the upper triangle in (6) to give an exact triangle in $D(\mathbb{Z}[G])$ of the form

(18)

$$Y_{K,S''}[-1]\oplus Y_{K,S''}[-2] \xrightarrow{\alpha} R\Gamma_c((\mathcal{O}_{K,S'})_{\mathcal{W}},\mathbb{Z}) \xrightarrow{\beta} R\Gamma_c((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{Z}) \longrightarrow .$$

After identifying the cohomology groups of the second and the third occurring complexes by using Proposition 2.4(iii) the long exact cohomology sequence of this triangle induces (after scaler extension) the sequence

$$\begin{aligned}0 &\longrightarrow \mathbb{Q}Y_{K,S''} \longrightarrow \mathbb{Q}Y_{K,S'}/\Delta_{S'}(\mathbb{Q}) \longrightarrow \mathbb{Q}Y_{K,S}/\Delta_S(\mathbb{Q})\\ &\xrightarrow{0} \mathbb{Q}Y_{K,S''} \xrightarrow{\mathrm{ord}^*_{S''}} \mathrm{Hom}_{\mathbb{Z}}(\mathcal{O}^{\times}_{K,S'},\mathbb{Q}) \xrightarrow{\pi_{S''}} \mathrm{Hom}_{\mathbb{Z}}(\mathcal{O}^{\times}_{K,S},\mathbb{Q}) \longrightarrow 0.\end{aligned}$$

Here $\mathrm{ord}^*_{S''}$ is induced by the linear dual of the map $\mathcal{O}^{\times}_{K,S'}\to Y_{K,S''}$ induced by taking valuations at each place in S''_K and $\pi_{S''}$ by the linear dual of the inclusion $\mathcal{O}_{K,S}^\times \subseteq \mathcal{O}_{K,S'}^\times$ and all other maps are obvious. This sequence implies that there is an exact commutative diagram

$$\begin{array}{ccccccccc} 0 & \to & \mathbb{R}Y_{K,S''} & \xrightarrow{H^1(\alpha)} & \mathbb{R}H^1_c((\mathcal{O}_{K,S'})_\mathcal{W},\mathbb{Z}) & \xrightarrow{H^1(\beta)} & \mathbb{R}H^1_c((\mathcal{O}_{K,S})_\mathcal{W},\mathbb{Z}) & \to & 0 \\ & & {\scriptstyle \eta_{S''}}\downarrow & & {\scriptstyle \lambda^*_{K,S'}}\downarrow & & {\scriptstyle \lambda^*_{K,S}}\downarrow & & \\ 0 & \to & \mathbb{R}Y_{K,S''} & \xrightarrow{H^2(\alpha)} & \mathbb{R}H^2_c((\mathcal{O}_{K,S'})_\mathcal{W},\mathbb{Z}) & \xrightarrow{H^2(\beta)} & \mathbb{R}H^2_c((\mathcal{O}_{K,S})_\mathcal{W},\mathbb{Z}) & \to & 0 \end{array}$$

where $\eta_{S''}$ sends each sum $\sum_{v\in S''}\sum_{w|v} x_w w$ to $\sum_{v\in S''}\sum_{w|v}\log(\mathrm{N}v)x_w w$. This diagram combines with the triangle (18) to imply that

$$\begin{aligned} &\vartheta_{\lambda^*_{K,S'}}(\det_G(R\Gamma_c((\mathcal{O}_{K,S'})_\mathcal{W},\mathbb{Z}))) \\ &= \det_{\mathbb{R}[G]}(\eta_{S''})^{-1}\vartheta_{\lambda^*_{K,S}}(\det_G(R\Gamma_c((\mathcal{O}_{K,S})_\mathcal{W},\mathbb{Z}))) \\ &= \Big(\prod_{v\in S''}\log(\mathrm{N}v)\Big)^{-1}\vartheta_{\lambda^*_{K,S}}(\det_G(R\Gamma_c((\mathcal{O}_{K,S})_\mathcal{W},\mathbb{Z}))). \end{aligned}$$

Since $\theta^*_{K/k,S'}(0) = \big(\prod_{v\in S''}\log(\mathrm{N}v)\big)\theta^*_{K/k,S}(0)$ this equality shows that (after changing S if necessary) we may assume that $\mathrm{Cl}_S(K)$ vanishes when verifying (17). Given this, the proposition follows from Proposition 2.9. □

3.3. Zeta elements. We now use the above results to reinterpret Conjecture 3.1 in terms of the existence of a canonical 'zeta element'. This interpretation will then play a key role in the proofs of Theorem 5.12, 5.16 and 7.5 given below.

The following definition of zeta element is in the same spirit as that used by Kato in [24] and [25].

Definition 3.5. The 'zeta element' $z_{K/k,S,T}$ of $\mathbb{G}_m$ relative to the data $K/k, S$ and T is the unique element of

$$\mathbb{R}\mathrm{det}_G(R\Gamma_T((\mathcal{O}_{K,S})_\mathcal{W},\mathbb{G}_m)) \cong \det_{\mathbb{R}[G]}(\mathbb{R}\mathcal{O}_{K,S}^\times)\otimes_{\mathbb{R}[G]}\det^{-1}_{\mathbb{R}[G]}(\mathbb{R}X_{K,S})$$

which satisfies $\vartheta_{\lambda_{K,S}}(z_{K/k,S,T}) = \theta^*_{K/k,S,T}(0)$.

The following 'leading term conjecture' is then our main object of study.

Conjecture 3.6 (LTC(K/k)). *In $\mathbb{R}\mathrm{det}_G(R\Gamma_T((\mathcal{O}_{K,S})_\mathcal{W},\mathbb{G}_m))$ one has an equality*

$$\mathbb{Z}[G]\cdot z_{K/k,S,T} = \det_G(R\Gamma_T((\mathcal{O}_{K,S})_\mathcal{W},\mathbb{G}_m)).$$

Given the definition of $z_{K/k,S,T}$, Proposition 3.4 implies immediately that this conjecture is equivalent to Conjecture 3.1 and hence is independent of the choices of S and T.

4. Preliminaries concerning exterior powers

In this section, we recall certain useful constructions concerning exterior powers and also prove algebraic results that are to be used in later sections.

4.1. Exterior powers. Let G be a finite abelian group. For a G-module M and $f \in \mathrm{Hom}_G(M, \mathbb{Z}[G])$, there is a G-homomorphism

$$\bigwedge_G^r M \longrightarrow \bigwedge_G^{r-1} M$$

for all $r \in \mathbb{Z}_{\geq 1}$, defined by

$$m_1 \wedge \cdots \wedge m_r \mapsto \sum_{i=1}^{r} (-1)^{i-1} f(m_i) m_1 \wedge \cdots \wedge m_{i-1} \wedge m_{i+1} \wedge \cdots \wedge m_r.$$

This morphism is also denoted by f.
This construction gives a homomorphism

$$\bigwedge_G^s \mathrm{Hom}_G(M, \mathbb{Z}[G]) \longrightarrow \mathrm{Hom}_G(\bigwedge_G^r M, \bigwedge_G^{r-s} M) \tag{19}$$

for all $r, s \in \mathbb{Z}_{\geq 0}$ such that $r \geq s$, defined by

$$f_1 \wedge \cdots \wedge f_s \mapsto (m \mapsto f_s \circ \cdots \circ f_1(m)).$$

By using this homomorphism we often regard an element of $\bigwedge_G^s \mathrm{Hom}_G(M, \mathbb{Z}[G])$ as an element of $\mathrm{Hom}_G(\bigwedge_G^r M, \bigwedge_G^{r-s} M)$.
For a G-algebra Q and a homomorphism f in $\mathrm{Hom}_G(M, Q)$, there is a G-homomorphism

$$\bigwedge_G^r M \longrightarrow (\bigwedge_G^{r-1} M) \otimes_G Q$$

defined by

$$m_1 \wedge \cdots \wedge m_r \mapsto \sum_{i=1}^{r} (-1)^{i-1} m_1 \wedge \cdots \wedge m_{i-1} \wedge m_{i+1} \wedge \cdots \wedge m_r \otimes f(m_i).$$

By the same method as the construction of (19), we have a homomorphism

$$\bigwedge_G^s \mathrm{Hom}_G(M, Q) \longrightarrow \mathrm{Hom}_G(\bigwedge_G^r M, (\bigwedge_G^{r-s} M) \otimes_G Q). \tag{20}$$

In the sequel we will find an explicit description of this homomorphism to be useful. This description is well-known and given by the following proposition, the proof of which we omit.

Proposition 4.1. *Let $m_1, \ldots, m_r \in M$ and $f_1, \ldots, f_s \in \mathrm{Hom}_G(M, Q)$. Then we have*

$$\begin{aligned}&(f_1 \wedge \cdots \wedge f_s)(m_1 \wedge \cdots \wedge m_r) \\ &= \sum_{\sigma \in \mathfrak{S}_{r,s}} \mathrm{sgn}(\sigma) m_{\sigma(s+1)} \wedge \cdots \wedge m_{\sigma(r)} \otimes \det(f_i(m_{\sigma(j)}))_{1 \leq i,j \leq s},\end{aligned}$$

where

$$\mathfrak{S}_{r,s} := \{\sigma \in \mathfrak{S}_r : \sigma(1) < \cdots < \sigma(s) \text{ and } \sigma(s+1) < \cdots < \sigma(r)\}.$$

In particular, if $r = s$, then we have

$$(f_1 \wedge \cdots \wedge f_r)(m_1 \wedge \cdots \wedge m_r) = \det(f_i(m_j))_{1\le i,j\le r}.$$

We will also find the technical observations that are contained in the next two results to be very useful.

Lemma 4.2. *Let E be a field and A an n-dimensional E-vector space. If we have an E-linear map*

$$\Psi : A \longrightarrow E^{\oplus m},$$

where $\Psi = \bigoplus_{i=1}^m \psi_i$ with $\psi_1, \ldots, \psi_m \in \mathrm{Hom}_E(A, E)$ $(m \le n)$, then we have

$$\mathrm{im}(\bigwedge_{1\le i\le m} \psi_i : \bigwedge_E^n A \longrightarrow \bigwedge_E^{n-m} A) = \begin{cases} \bigwedge_E^{n-m} \ker(\Psi), & \text{if } \Psi \text{ is surjective,} \\ 0, & \text{if } \Psi \text{ is not surjective.} \end{cases}$$

Proof. Suppose first that Ψ is surjective. Then there exists a subspace $B \subset A$ such that $A = \ker(\Psi) \oplus B$ and Ψ maps B isomorphically onto $E^{\oplus m}$. We see that $\bigwedge_{1\le i\le m} \psi_i$ induces an isomorphism

$$\bigwedge_E^m B \xrightarrow{\sim} E.$$

Hence we have an isomorphism

$$\bigwedge_{1\le i\le m} \psi_i : \bigwedge_E^n A = \bigwedge_E^{n-m} \ker(\Psi) \otimes_E \bigwedge_E^m B \xrightarrow{\sim} \bigwedge_E^{n-m} \ker(\Psi).$$

In particular, we have

$$\mathrm{im}(\bigwedge_{1\le i\le m} \psi_i : \bigwedge_E^n A \longrightarrow \bigwedge_E^{n-m} A) = \bigwedge_E^{n-m} \ker(\Psi).$$

Next, suppose that Ψ is not surjective. Then $\psi_1, \ldots, \psi_m \in \mathrm{Hom}_E(A, E)$ are linearly dependent. In fact, since each ψ_i is contained in $\mathrm{Hom}_E(A/\ker(\Psi), E)$, we have

$$\dim_E(\langle \psi_1, \ldots, \psi_m \rangle) \le \dim_E(A/\ker(\Psi)) = \dim_E(\mathrm{im}(\Psi)),$$

so $\dim_E(\langle \psi_1, \ldots, \psi_m \rangle) < m$ if $\dim_E(\mathrm{im}(\Psi)) < m$. This shows that the element $\bigwedge_{1\le i\le m} \psi_i$ vanishes, as required. □

Using the same notation as in Lemma 4.2, we now consider an endomorphism $\psi \in \mathrm{End}_E(A)$. We write r_ψ for the dimension over E of $\ker(\psi)$ and consider the composite isomorphism

$$\begin{aligned} F_\psi : \bigwedge_E^n A \otimes_E \bigwedge_E^n \mathrm{Hom}_E(A, E) \quad &\simeq \quad \det_E(A) \otimes_E \det_E^{-1}(A) \\ &\xrightarrow{\sim} \quad \det_E(\ker(\psi)) \otimes_E \det_E^{-1}(\mathrm{coker}(\psi)) \\ &\simeq \quad \bigwedge_E^{r_\psi} \ker(\psi) \otimes_E \bigwedge_E^{r_\psi} \mathrm{Hom}_E(\mathrm{coker}(\psi), E), \end{aligned}$$

where the second isomorphism is induced by the tautological exact sequence

$$0 \longrightarrow \ker(\psi) \longrightarrow A \xrightarrow{\psi} A \longrightarrow \mathrm{coker}(\psi) \longrightarrow 0.$$

Then the proof of Lemma 4.2 leads directly to the following useful description of this isomorphism F_ψ.

Lemma 4.3. *With E, A and ψ as above, we fix an E-basis $\{b_1, \ldots, b_n\}$ of A so that $\mathrm{im}(\psi) = \langle b_{r_\psi+1}, \ldots, b_n\rangle$ and write $\{b_1^*, \ldots, b_n^*\}$ for the corresponding dual basis of $\mathrm{Hom}_E(A, E)$. For each index i we also set $\psi_i := b_i^* \circ \psi$.*
Then for every a in $\bigwedge_E^n A$ the element $(\bigwedge_{r_\psi < i \le n} \psi_i)(a)$ belongs to $\bigwedge_E^{r_\psi} \ker(\psi)$ and one has

$$F_\psi(a \otimes (b_1^* \wedge \cdots \wedge b_n^*)) = (-1)^{r_\psi(n-r_\psi)}(\bigwedge_{r_\psi < i \le n} \psi_i)(a) \otimes (b_1^* \wedge \cdots \wedge b_{r_\psi}^*).$$

Here, on the right hand side of the equation, we use the equality $\mathrm{im}(\psi) = \langle b_{r_\psi+1}, \ldots, b_n\rangle$ to regard b_i^* for each i with $1 \le i \le r_\psi$ as an element of $\mathrm{Hom}_E(\mathrm{coker}(\psi), E)$.

4.2. Rubin lattices. The following definition is due to Rubin [45, §1.2]. We adopt the notation in [46] for the lattice. Note in particular that the notation '$\bigcap$' does not refer to an intersection.

Definition 4.4. For a finitely generated G-module M and a non-negative integer r we define the 'r-th Rubin lattice' by setting

$$\bigcap_G^r M = \{m \in \mathbb{Q}\bigwedge_G^r M : \Phi(m) \in \mathbb{Z}[G] \text{ for all } \Phi \in \bigwedge_G^r \mathrm{Hom}_G(M, \mathbb{Z}[G])\}.$$

In particular, one has $\bigcap_G^0 M = \mathbb{Z}[G]$.

Remark 4.5. We define the homomorphism $\iota : \bigwedge_G^r \mathrm{Hom}_G(M, \mathbb{Z}[G]) \to \mathrm{Hom}_G(\bigwedge_G^r M, \mathbb{Z}[G])$ by sending each element $\varphi_1 \wedge \cdots \wedge \varphi_r$ to $\varphi_r \circ \cdots \circ \varphi_1$ (see (19)). Then it is not difficult to see that the map

$$\bigcap_G^r M \xrightarrow{\sim} \mathrm{Hom}_G(\mathrm{im}(\iota), \mathbb{Z}[G]);\ m \mapsto (\Phi \mapsto \Phi(m))$$

is an isomorphism (see [45, §1.2]).

By this remark, one obtains the following result.

Proposition 4.6. *Let P be a finitely generated projective G-module. Then we have*

$$\bigcap_G^r P = \bigwedge_G^r P$$

for all non-negative integers r.

Lemma 4.7. *Let M be a G-module. Suppose that there is a finitely generated projective G-module P and an injection $j : M \hookrightarrow P$ whose cokernel is $\mathbb{Z}$-torsion-free.*

(i) *The map*

$$\mathrm{Hom}_G(P, \mathbb{Z}[G]) \longrightarrow \mathrm{Hom}_G(M, \mathbb{Z}[G])$$

induced by j is surjective.

(ii) *If we regard M as a submodule of P via j, then we have*

$$\bigcap_G^r M = (\mathbb{Q}\bigwedge_G^r M) \cap \bigwedge_G^r P.$$

Proof. The assertion (i) follows from [45, Prop. 1.1 (ii)]. Note that

$$\bigwedge_G^r \mathrm{Hom}_G(P, \mathbb{Z}[G]) \longrightarrow \bigwedge_G^r \mathrm{Hom}_G(M, \mathbb{Z}[G])$$

is also surjective. This induces a surjection

$$\mathrm{im}(\iota_P) \longrightarrow \mathrm{im}(\iota_M),$$

where ι_P and ι_M denote the maps defined in Remark 4.5 for P and M, respectively. Hence, taking the dual, we have an injection

$$\bigcap_G^r M \simeq \mathrm{Hom}_G(\mathrm{im}(\iota_M), \mathbb{Z}[G]) \longrightarrow \mathrm{Hom}_G(\mathrm{im}(\iota_P), \mathbb{Z}[G]) \simeq \bigcap_G^r P.$$

Since P is projective, we have $\bigcap_G^r P = \bigwedge_G^r P$ by Proposition 4.6. Hence we have

$$\bigcap_G^r M \subset \bigwedge_G^r P.$$

Next, we show the reverse inclusion '$\supset$'. To do this we fix a in $(\mathbb{Q}\bigwedge_G^r M) \cap \bigwedge_G^r P$ and Φ in $\bigwedge_G^r \mathrm{Hom}_G(M, \mathbb{Z}[G])$. By (i), we can take a lift $\widetilde{\Phi} \in \bigwedge_G^r \mathrm{Hom}_G(P, \mathbb{Z}[G])$ of Φ. Since $a \in \bigwedge_G^r P$, we have

$$\Phi(a) = \widetilde{\Phi}(a) \in \mathbb{Z}[G].$$

This shows that a belongs to $\bigcap_G^r M$, as required. $\square$

Remark 4.8. The proof of Lemma 4.7 shows that the cokernel of the injection

$$\bigcap_G^r M \longrightarrow \bigwedge_G^r P$$

is $\mathbb{Z}$-torsion-free. This implies that for any abelian group A, the map

$$(\bigcap_G^r M) \otimes_{\mathbb{Z}} A \longrightarrow (\bigwedge_G^r P) \otimes_{\mathbb{Z}} A$$

is injective.

4.3. Homomorphisms between Rubin lattices. In the sequel we fix a subgroup H of G and an ideal J of $\mathbb{Z}[H]$. Recall that we denote the augmentation ideal of $\mathbb{Z}[H]$ by $I(H)$. Put $J_H := J/I(H)J$. We also put $\mathcal{J} := J\mathbb{Z}[G]$, and $\mathcal{J}_H := \mathcal{J}/I(H)\mathcal{J}$.

Proposition 4.9. *We have a natural isomorphism of G/H-modules*

$$\mathcal{J}_H \simeq \mathbb{Z}[G/H] \otimes_{\mathbb{Z}} J_H.$$

Proof. Define a homomorphism

$$\mathbb{Z}[G/H] \otimes_{\mathbb{Z}} J_H \longrightarrow \mathcal{J}_H$$

by $\tau \otimes \overline{a} \mapsto \overline{\widetilde{\tau}a}$, where $\tau \in G/H$, $a \in J$, and $\widetilde{\tau} \in G$ is a lift of τ. One can easily check that this homomorphism is well-defined, and bijective. □

Definition 4.10. Let M be a G-module. For $\varphi \in \mathrm{Hom}_G(M, \mathbb{Z}[G])$, we define $\varphi^H \in \mathrm{Hom}_{G/H}(M^H, \mathbb{Z}[G/H])$ by

$$M^H \stackrel{\varphi}{\longrightarrow} \mathbb{Z}[G]^H \simeq \mathbb{Z}[G/H],$$

where the last isomorphism is given by $\mathrm{N}_H = \sum_{\sigma\in H} \sigma \mapsto 1$. Let r be a non-negative integer. For $\Phi \in \bigwedge_G^r \mathrm{Hom}_G(M, \mathbb{Z}[G])$, we define $\Phi^H \in \bigwedge_{G/H}^r \mathrm{Hom}_{G/H}(M^H, \mathbb{Z}[G/H])$ to be the image of Φ under the map

$$\varphi_1 \wedge \cdots \wedge \varphi_r \mapsto \varphi_1^H \wedge \cdots \wedge \varphi_r^H.$$

For convention, if $r = 0$, then we define $\Phi^H \in \mathbb{Z}[G/H]$ to be the image of $\Phi \in \mathbb{Z}[G]$ under the natural map $: \mathbb{Z}[G] \longrightarrow \mathbb{Z}[G/H]$.

Proposition 4.11. *Let M be a G-module and $r \in \mathbb{Z}_{\geq 0}$. For any $m \in \mathbb{Q}\bigwedge_G^r M$ and $\Phi \in \bigwedge_G^r \mathrm{Hom}_G(M, \mathbb{Z}[G])$, we have*

$$\Phi(m) = \Phi^H(\mathrm{N}_H^r m) \text{ in } \mathbb{Q}[G/H],$$

where N_H^r denote the map $\mathbb{Q}\bigwedge_G^r M \to \mathbb{Q}\bigwedge_{G/H}^r M^H$ induced by $\mathrm{N}_H : M \to M^H$.

Proof. This follows directly from the definition of Φ^H. □

We consider the canonical map

$$\nu : \bigcap_{G/H}^{r} M^H \longrightarrow \bigcap_{G}^{r} M$$

which is defined as follows. Let

$$\iota : \bigwedge_G^r \mathrm{Hom}_G(M, \mathbb{Z}[G]) \longrightarrow \mathrm{Hom}_G(\bigwedge_G^r M, \mathbb{Z}[G])$$

and

$$\iota_H : \bigwedge_{G/H}^r \mathrm{Hom}_{G/H}(M^H, \mathbb{Z}[G/H]) \longrightarrow \mathrm{Hom}_{G/H}(\bigwedge_{G/H}^r M^H, \mathbb{Z}[G/H])$$

be the homomorphisms defined in Remark 4.5. The map

$$\mathrm{im}(\iota) \longrightarrow \mathrm{im}(\iota_H);\ \iota(\Phi) \mapsto \iota_H(\Phi^H)$$

induces a map

$$\alpha : \mathrm{Hom}_G(\mathrm{im}(\iota_H), \mathbb{Z}[G]) \longrightarrow \mathrm{Hom}_G(\mathrm{im}(\iota), \mathbb{Z}[G]) \simeq \bigcap_G^r M.$$

Note that we have a canonical isomorphism

$$\beta : \mathrm{Hom}_G(\mathrm{im}(\iota_H), \mathbb{Z}[G]) \xrightarrow{\sim} \mathrm{Hom}_{G/H}(\mathrm{im}(\iota_H), \mathbb{Z}[G/H]) \simeq \bigcap_{G/H}^r M^H;\ \varphi \mapsto \varphi^H.$$

We define a map ν by

$$\nu := \alpha \circ \beta^{-1} : \bigcap_{G/H}^r M^H \longrightarrow \bigcap_G^r M.$$

Proposition 4.12. *Let M be a finitely generated G-module which is $\mathbb{Z}$-torsion-free. For any $r \in \mathbb{Z}_{\geq 0}$, the map $\nu : \bigcap_{G/H}^r M^H \to \bigcap_G^r M$ is injective. Furthermore, the maps*

$$(\bigcap_{G/H}^r M^H) \otimes_{\mathbb{Z}} J_H \longrightarrow (\bigcap_G^r M) \otimes_{\mathbb{Z}} J_H \longrightarrow (\bigcap_G^r M) \otimes_{\mathbb{Z}} \mathbb{Z}[H]/I(H)J$$

are both injective, where the first map is induced by ν, and the second by inclusion $J_H \hookrightarrow \mathbb{Z}[H]/I(H)J$.

Proof. The proof is the same as [46, Lem. 2.11], so we omit it. □

Remark 4.13. The inclusion $M^H \subset M$ induces a map

$$\xi : \bigcap_{G/H}^r M^H \longrightarrow \bigcap_G^r M.$$

We note that this map *does not coincide* with the above map ν if $r > 1$. Indeed, one can check that $\mathrm{im}(\xi) \subset |H|^{\max\{0,r-1\}} \bigcap_G^r M$ (see [37, Lem. 4.8]), and

$$\nu = |H|^{-\max\{0,r-1\}}\xi.$$

Remark 4.14. Let P be a finitely generated projective G-module. Then, any element of P^H is written as $\mathrm{N}_H\, a$ with some $a \in P$, since P is cohomologically trivial. One can check that, if $r > 0$ (resp. $r = 0$), then the map $\nu : \bigwedge_{G/H}^r P^H \to \bigwedge_G^r P$ constructed above coincides with the map

$$\mathrm{N}_H\, a_1 \wedge \cdots \wedge \mathrm{N}_H\, a_r \mapsto \mathrm{N}_H\, a_1 \wedge \cdots \wedge a_r$$

$$(\text{resp. } \mathbb{Z}[G/H] \simeq \mathbb{Z}[G]^H \hookrightarrow \mathbb{Z}[G]).$$

In particular, we know that $\mathrm{im}(\nu) = \mathrm{N}_H \bigwedge_G^r P$.

Proposition 4.15. *Let M be a finitely generated G-module which is $\mathbb{Z}$-torsion-free, and $r \in \mathbb{Z}_{\geq 0}$. Then the map*

$$(\bigcap_{G/H}^{r} M^H) \otimes_{\mathbb{Z}} J_H \longrightarrow \mathrm{Hom}_G(\bigwedge_G^r \mathrm{Hom}_G(M, \mathbb{Z}[G]), \mathcal{J}_H);\ \alpha \mapsto (\Phi \mapsto \Phi^H(\alpha))$$

is injective. (We regard $\Phi^H(\alpha) \in \mathbb{Z}[G/H] \otimes_{\mathbb{Z}} J_H$ as an element of $\mathcal{J}_H$ via the isomorphism $\mathcal{J}_H \simeq \mathbb{Z}[G/H] \otimes_{\mathbb{Z}} J_H$ in Proposition 4.9.)

Proof. The proof is the same as [46, Th. 2.12]. □

4.4. Congruences between exterior powers. The following definition is originally due to Darmon [12], and used in [46, Def. 2.13] and [37, Def. 5.1].

Definition 4.16. Let M be a G-module. For $m \in M$, define

$$\mathcal{N}_H(m) = \sum_{\sigma \in H} \sigma m \otimes \sigma^{-1} \in M \otimes_{\mathbb{Z}} \mathbb{Z}[H]/I(H)J.$$

The following proposition is an improvement of the result of the third author in [46, Prop. 2.15].

Proposition 4.17. *Let P be a finitely generated projective G-module, $r \in \mathbb{Z}_{\geq 0}$, and*

$$\nu : (\bigwedge_{G/H}^{r} P^H) \otimes_{\mathbb{Z}} J_H \longrightarrow (\bigwedge_G^r P) \otimes_{\mathbb{Z}} \mathbb{Z}[H]/I(H)J$$

be the injection in Proposition 4.12. For an element $a \in \bigwedge_G^r P$, the following are equivalent.

(i) $a \in \mathcal{J} \cdot \bigwedge_G^r P$,
(ii) $\mathcal{N}_H(a) \in \mathrm{im}(\nu)$,
(iii) $\Phi(a) \in \mathcal{J}$ *for every* $\Phi \in \bigwedge_G^r \mathrm{Hom}_G(P, \mathbb{Z}[G])$.

Furthermore, if the above equivalent conditions are satisfied, then for every $\Phi \in \bigwedge_G^r \mathrm{Hom}_G(P, \mathbb{Z}[G])$ we have

$$\Phi(a) = \Phi^H(\nu^{-1}(\mathcal{N}_H(a))) \text{ in } \mathcal{J}_H,$$

where we regard $\Phi^H(\nu^{-1}(\mathcal{N}_H(a))) \in \mathbb{Z}[G/H] \otimes_{\mathbb{Z}} J_H$ as an element of $\mathcal{J}_H$ via the isomorphism $\mathcal{J}_H \simeq \mathbb{Z}[G/H] \otimes_{\mathbb{Z}} J_H$ in Proposition 4.9.

Proof. By Swan's Theorem (see [11, (32.1)]), for every prime p, P_p is a free $\mathbb{Z}_p[G]$-module of rank, d say, independent of p. Considering locally, we may assume that P is a free G-module of rank d. We may assume $r \leq d$. Clearly, (i) implies (iii). We shall show that (iii) implies (ii). Suppose $\Phi(a) \in \mathcal{J}$ for all $\Phi \in \bigwedge_G^r \mathrm{Hom}_G(P, \mathbb{Z}[G])$. Fix a basis $\{b_1, \ldots, b_d\}$ of P. Write

$$a = \sum_{\mu \in \mathfrak{S}_{d,r}} x_\mu b_{\mu(1)} \wedge \cdots \wedge b_{\mu(r)},$$

with some $x_\mu \in \mathbb{Z}[G]$. For each μ, by Proposition 4.1, we have

$$x_\mu = (b_{\mu(1)}^* \wedge \cdots \wedge b_{\mu(r)}^*)(a) \in \mathcal{J},$$

where $b_i^* \in \mathrm{Hom}_G(P, \mathbb{Z}[G])$ is the dual basis of b_i. For each $\tau \in G/H$, fix a lift $\widetilde{\tau} \in G$. Note that we have a direct sum decomposition

$$\mathcal{J} = \bigoplus_{\tau \in G/H} J\widetilde{\tau}.$$

Therefore, we can write each x_μ as follows:

$$x_\mu = \sum_{\tau \in G/H} y_{\tau\mu}\widetilde{\tau},$$

where $y_{\tau\mu} \in J$. Hence we have

$$\begin{aligned}
\mathcal{N}_H(a) &= \sum_{\sigma \in H} \sum_{\mu \in \mathfrak{S}_{d,r}} \sum_{\tau \in G/H} \sigma y_{\tau\mu}\widetilde{\tau} b_{\mu(1)} \wedge \cdots \wedge b_{\mu(r)} \otimes \sigma^{-1} \\
&= \sum_{\sigma \in H} \sum_{\mu \in \mathfrak{S}_{d,r}} \sum_{\tau \in G/H} \sigma\widetilde{\tau} b_{\mu(1)} \wedge \cdots \wedge b_{\mu(r)} \otimes \sigma^{-1} y_{\tau\mu} \\
&= \sum_{\mu \in \mathfrak{S}_{d,r}} \sum_{\tau \in G/H} \mathrm{N}_H\, \widetilde{\tau} b_{\mu(1)} \wedge \cdots \wedge b_{\mu(r)} \otimes y_{\tau\mu} \\
&\in \mathrm{N}_H \bigwedge_G^r P \otimes_{\mathbb{Z}} J_H = \mathrm{im}(\nu)
\end{aligned}$$

(see Remark 4.14). This shows (ii). We also see by Remark 4.14 that

$$\nu^{-1}(\mathcal{N}_H(a)) = \sum_{\mu \in \mathfrak{S}_{d,r}} \sum_{\tau \in G/H} \tau\, \mathrm{N}_H\, b_{\mu(1)} \wedge \cdots \wedge \mathrm{N}_H\, b_{\mu(r)} \otimes y_{\tau\mu} \in (\bigwedge_{G/H}^r P^H) \otimes_{\mathbb{Z}} J_H.$$

Hence, by Proposition 4.11, we have

$$\Phi(a) = \Phi^H(\nu^{-1}(\mathcal{N}_H(a))) \quad in \quad \mathcal{J}_H$$

for all $\Phi \in \bigwedge_G^r \mathrm{Hom}_G(P, \mathbb{Z}[G])$.
Finally, we show that (ii) implies (i). Suppose $\mathcal{N}_H(a) \in \mathrm{im}(\nu) = (\mathrm{N}_H \bigwedge_G^r P) \otimes_{\mathbb{Z}} J_H$. As before, we write

$$a = \sum_{\mu \in \mathfrak{S}_{d,r}} \sum_{\tau \in G/H} y_{\tau\mu}\widetilde{\tau} b_{\mu(1)} \wedge \cdots \wedge b_{\mu(r)}$$

with $y_{\tau\mu} \in \mathbb{Z}[H]$. We have

$$\mathcal{N}_H(a) = \sum_{\sigma \in H} \sum_{\mu \in \mathfrak{S}_{d,r}} \sum_{\tau \in G/H} \sigma\widetilde{\tau} b_{\mu(1)} \wedge \cdots \wedge b_{\mu(r)} \otimes \sigma^{-1} y_{\tau\mu} \in (\mathrm{N}_H \bigwedge_G^r P) \otimes_{\mathbb{Z}} J_H.$$

Since $(\bigwedge_G^r P) \otimes_{\mathbb{Z}} \mathbb{Z}[H]/I(H)J \simeq \bigoplus_{\sigma,\mu,\tau} \mathbb{Z}[H]/I(H)J$ as abelian groups, we must have $y_{\tau\mu} \in J$. This shows that $a \in \mathcal{J} \cdot \bigwedge_G^r P$. $\square$

5. Congruences for Rubin-Stark elements

For a finite abelian extension K/k, and an intermediate field L, a conjecture which describes a congruence relation between two Rubin-Stark elements for K/k and L/k was formulated by the third author in [46, Conj. 3]. Mazur and Rubin also formulated in [37, Conj. 5.2] essentially the same conjecture. In this section, we formulate a refined version (see Conjecture 5.4) of these conjectures. We also recall a conjecture formulated by the first author, which was studied in [22], [3], [16], [17], [48], and [46] (see Conjecture 5.9). In [46, Th. 3.15], the third author proved a link between Conjecture 5.4 and Conjecture 5.9. We now improve the argument given there to show that Conjecture 5.4 and Conjecture 5.9 are in fact equivalent (see Theorem 5.10). Finally we prove that the natural equivariant leading term conjecture (Conjecture 3.6) implies both Conjecture 5.4 and Conjecture 5.9 (see Theorem 5.16).

5.1. The Rubin-Stark conjecture. In this subsection, we recall the formulation of the Rubin-Stark conjecture [45, Conj. B′].

Let $K/k, G, S, T$ be as in §3, namely, K/k is a finite abelian extension of global fields, G is its Galois group, S is a non-empty finite set of places of k such that $S_\infty(k) \cup S_{\rm ram}(K/k) \subset S$, and T is a finite set of places of k which is disjoint from S. In this section, we assume that $\mathcal{O}_{K,S,T}^\times$ is $\mathbb{Z}$-torsion-free.

Following Rubin [45, Hyp. 2.1] we assume that S satisfies the following hypothesis with respect to some chosen integer r with $0 \le r < |S|$: there exists a subset $V \subset S$ of order r such that each place in V splits completely in K/k.

Recall that for any $\chi \in \widehat{G}$ we denote by $r_{\chi,S}$ the order of vanishing of $L_{k,S,T}(\chi, s)$ at $s = 0$. We know by [50, Chap. I, Prop. 3.4] that

$$(21)\quad r_{\chi,S} = \dim_{\mathbb{C}}(e_\chi \mathbb{C}X_{K,S}) = \begin{cases} |\{v \in S : \chi(G_v) = 1\}| & \text{if } \chi \neq 1, \\ |S| - 1 & \text{if } \chi = 1. \end{cases}$$

Therefore, the existence of V ensures that $r \le r_{\chi,S}$ for every χ and hence the function $s^{-r}L_{k,S,T}(\chi, s)$ is holomorphic at $s = 0$. We define the 'r-th order Stickelberger element' by

$$\theta_{K/k,S,T}^{(r)} := \lim_{s\to 0} \sum_{\chi \in \widehat{G}} s^{-r} L_{k,S,T}(\chi^{-1}, s) e_\chi \in \mathbb{R}[G].$$

Note that the 0-th order Stickelberger element $\theta_{K/k,S,T}^{(0)} (= \theta_{K/k,S,T}(0))$ is the usual Stickelberger element.

Recall that we have the regulator isomorphism

$$\lambda_{K,S} : \mathbb{R}\mathcal{O}_{K,S,T}^\times \xrightarrow{\sim} \mathbb{R}X_{K,S}$$

defined by

$$\lambda_{K,S}(a) = -\sum_{w \in S_K} \log|a|_w w.$$

This map $\lambda_{K,S}$ induces the isomorphism

$$\bigwedge_{\mathbb{R}[G]}^{r} \mathbb{R}\mathcal{O}_{K,S,T}^{\times} \xrightarrow{\sim} \bigwedge_{\mathbb{R}[G]}^{r} \mathbb{R}X_{K,S},$$

which we also denote by $\lambda_{K,S}$. For each place $v \in S$, fix a place w of K lying above v. Take any $v_0 \in S \setminus V$, and define the '(r-th order) Rubin-Stark element'

$$\epsilon_{K/k,S,T}^{V} \in \bigwedge_{\mathbb{R}[G]}^{r} \mathbb{R}\mathcal{O}_{K,S,T}^{\times} = \mathbb{R}\bigwedge_{G}^{r} \mathcal{O}_{K,S,T}^{\times}.$$

by

$$\lambda_{K,S}(\epsilon_{K/k,S,T}^{V}) = \theta_{K/k,S,T}^{(r)} \bigwedge_{v \in V} (w - w_0), \tag{22}$$

where $\bigwedge_{v \in V}(w - w_0)$ is arranged by some chosen order of the elements in V. One can show that the Rubin-Stark element $\epsilon_{K/k,S,T}^{V}$ does not depend on the choice of $v_0 \in S \setminus V$.

We consider the Rubin lattice

$$\bigcap_{G}^{r} \mathcal{O}_{K,S,T}^{\times} \subset \mathbb{Q}\bigwedge_{G}^{r} \mathcal{O}_{K,S,T}^{\times}$$

(see Definition 4.4). The Rubin-Stark conjecture claims

Conjecture 5.1 (The Rubin-Stark conjecture for $(K/k, S, T, V)$). *One has*

$$\epsilon_{K/k,S,T}^{V} \in \bigcap_{G}^{r} \mathcal{O}_{K,S,T}^{\times}.$$

Remark 5.2. One can check that the above Rubin-Stark conjecture is equivalent to [45, Conj. B′] for the data $(K/k, S, T, V)$, and that our Rubin-Stark element $\epsilon_{K/k,S,T}^{V}$ coincides with the unique element predicted by [45, Conj. B′]. This shows, in particular, that the validity of the conjecture does not depend on the choice of the places lying above $v \in S$ or on the ordering of the elements in V.

Remark 5.3. The Rubin-Stark conjecture for $(K/k, S, T, V)$ is known to be true in the following cases:

- (i) $r = 0$. In this case $\epsilon_{K/k,S,T}^{\emptyset} = \theta_{K/k,S,T}^{(0)} = \theta_{K/k,S,T}(0) \in \mathbb{R}[G]$ so the Rubin-Stark conjecture claims only that $\theta_{K/k,S,T}(0) \in \mathbb{Z}[G]$ which is a celebrated result of Deligne-Ribet, Cassou-Noguès, and Barsky.
- (ii) $[K : k] \leq 2$. This is due to Rubin [45, Cor. 3.2 and Th. 3.5].
- (iii) K is an abelian extension over $\mathbb{Q}$. This is due to the first author [3, Th. A].
- (iv) k is a global function field. This is due to the first author [3, Th. A].

5.2. Conventions for Rubin-Stark elements. The notation $\epsilon^V_{K/k,S,T}$ has some ambiguities, since $\epsilon^V_{K/k,S,T}$ depends on the choice of the places lying above $v \in S$, and on the choice of the order of the elements in V. To avoid this ambiguity, we use the following convention: when we consider the Rubin-Stark element $\epsilon^V_{K/k,S,T}$, we always fix a place w of K lying above each $v \in S$, and label the elements of S as

$$S = \{v_0, v_1, \dots, v_n\}$$

so that $V = \{v_1, \dots, v_r\}$, and thus we fix the order of the elements in V. So, under this convention, the Rubin-Stark element $\epsilon^V_{K/k,S,T}$ is the element characterized by

$$\lambda_{K,S}(\epsilon^V_{K/k,S,T}) = \theta^{(r)}_{K/k,S,T} \bigwedge_{1\le i\le r} (w_i - w_0).$$

5.3. Conjectures on Rubin-Stark elements. In this subsection, we give a refinement of the conjecture formulated by the third author [46, Conj. 3], and Mazur and Rubin [37, Conj. 5.2]. Let $K/k, G, S, T$ be as before, and we assume that, for a non-negative integer r, there exists a subset $V \subset S$ of order r such that each place in V splits completely in K. We fix a subgroup H of G for which, for some integer r' with $r' \ge r$, there exists a subset $V' \subset S$ of order r', which contains V, and satisfies that each place in V' splits completely in the field $L := K^H$.

Following the convention in §5.2, we fix, for each place $v \in S$, a place w of K lying above v, and label the elements of S as $S = \{v_0, \dots, v_n\}$ so that $V = \{v_1, \dots, v_r\}$ and $V' = \{v_1, \dots, v_{r'}\}$. We consider the Rubin-Stark elements $\epsilon^V_{K/k,S,T}$ and $\epsilon^{V'}_{L/k,S,T}$ characterized by

$$\lambda_{K,S}(\epsilon^V_{K/k,S,T}) = \theta^{(r)}_{K/k,S,T} \bigwedge_{1\le i\le r} (w_i - w_0)$$

and

$$\lambda_{L,S}(\epsilon^{V'}_{L/k,S,T}) = \theta^{(r')}_{L/k,S,T} \bigwedge_{1\le i\le r'} (w_i - w_0)$$

respectively, where we denote the place of L lying under w also by w.

For each integer i with $1 \le i \le n$ we write G_i for the decomposition group of v_i in G. For any subgroup $U \subset G$, recall that the augmentation ideal of $\mathbb{Z}[U]$ is denoted by $I(U)$. Put $\mathcal{I}_i := I(G_i)\mathbb{Z}[G]$ and $I_H := I(H)\mathbb{Z}[G]$. We define

$$\mathrm{Rec}_i : \mathcal{O}^\times_{L,S,T} \longrightarrow (\mathcal{I}_i)_H = \mathcal{I}_i / I_H \mathcal{I}_i$$

by

$$\mathrm{Rec}_i(a) = \sum_{\tau \in G/H} \tau^{-1}(\mathrm{rec}_{w_i}(\tau a) - 1).$$

Here, rec_{w_i} is the reciprocity map $L^\times_{w_i} \to G_i$ at w_i. Note that $\tau^{-1}(\mathrm{rec}_{w_i}(\tau a) - 1)$ is well-defined for $\tau \in G/H$ in $(\mathcal{I}_i)_H$.

We put $W := V' \setminus V = \{v_{r+1}, \ldots, v_{r'}\}$. We define an ideal J_W of $\mathbb{Z}[H]$ by

$$J_W := \begin{cases} (\prod_{r<i\le r'} I(G_i))\mathbb{Z}[H], & \text{if } W \neq \emptyset, \\ \mathbb{Z}[H], & \text{if } W = \emptyset, \end{cases}$$

and put $(J_W)_H := J_W/I(H)J_W$. We also define an ideal $\mathcal{J}_W$ of $\mathbb{Z}[G]$ by

$$\mathcal{J}_W := \begin{cases} \prod_{r<i\le r'} \mathcal{I}_i, & \text{if } W \neq \emptyset, \\ \mathbb{Z}[G], & \text{if } W = \emptyset, \end{cases}$$

and put $(\mathcal{J}_W)_H := \mathcal{J}_W/I_H\mathcal{J}_W$. Note that $\mathcal{J}_W = J_W\mathbb{Z}[G]$. By Proposition 4.9, we have a natural isomorphism of G/H-modules $\mathbb{Z}[G/H] \otimes_{\mathbb{Z}} (J_W)_H \simeq (\mathcal{J}_W)_H$. We consider the graded G/H-algebra

$$\mathcal{Q}_W := \bigoplus_{a_1,\ldots,a_{r'-r}\in\mathbb{Z}_{\ge 0}} (\mathcal{I}_{r+1}^{a_1}\cdots\mathcal{I}_{r'}^{a_{r'-r}})_H,$$

where

$$(\mathcal{I}_{r+1}^{a_1}\cdots\mathcal{I}_{r'}^{a_{r'-r}})_H := \mathcal{I}_{r+1}^{a_1}\cdots\mathcal{I}_{r'}^{a_{r'-r}}/I_H\mathcal{I}_{r+1}^{a_1}\cdots\mathcal{I}_{r'}^{a_{r'-r}},$$

and we define the 0-th power of any ideal of $\mathbb{Z}[G]$ to be $\mathbb{Z}[G]$.

For any integer i with $r < i \le r'$ we regard Rec_i as an element of $\mathrm{Hom}_{G/H}(\mathcal{O}_{L,S,T}^\times, \mathcal{Q}_W)$ via the natural embedding $(\mathcal{I}_i)_H \hookrightarrow \mathcal{Q}_W$. Then by the same method as in [46, Prop. 2.7] (or [37, Cor. 2.1]), one shows that $\bigwedge_{r<i\le r'} \mathrm{Rec}_i \in \bigwedge_{G/H}^{r'-r} \mathrm{Hom}_{G/H}(\mathcal{O}_{L,S,T}^\times, \mathcal{Q}_W)$ induces the map

$$\bigcap_{G/H}^{r'} \mathcal{O}_{L,S,T}^\times \longrightarrow (\bigcap_{G/H}^{r} \mathcal{O}_{L,S,T}^\times) \otimes_{G/H} (\mathcal{J}_W)_H \simeq (\bigcap_{G/H}^{r} \mathcal{O}_{L,S,T}^\times) \otimes_{\mathbb{Z}} (J_W)_H, \tag{23}$$

which we denote by Rec_W.

Following Definition 4.16, we define

$$\mathcal{N}_H : \bigcap_G^r \mathcal{O}_{K,S,T}^\times \longrightarrow (\bigcap_G^r \mathcal{O}_{K,S,T}^\times) \otimes_{\mathbb{Z}} \mathbb{Z}[H]/I(H)J_W$$

by $\mathcal{N}_H(a) = \sum_{\sigma\in H} \sigma a \otimes \sigma^{-1}$.

Note that since $(\mathcal{O}_{K,S,T}^\times)^H = \mathcal{O}_{L,S,T}^\times$, there is a natural injective homomorphism

$$\nu : (\bigcap_{G/H}^r \mathcal{O}_{L,S,T}^\times) \otimes_{\mathbb{Z}} (J_W)_H \longrightarrow (\bigcap_G^r \mathcal{O}_{K,S,T}^\times) \otimes_{\mathbb{Z}} \mathbb{Z}[H]/I(H)J_W$$

by Proposition 4.12.

To state the following conjecture we assume the validity of the Rubin-Stark conjecture (Conjecture 5.1) for both $(K/k, S, T, V)$ and $(L/k, S, T, V')$.

Conjecture 5.4 ($\mathrm{MRS}(K/L/k, S, T, V, V')$). *The element $\mathcal{N}_H(\epsilon_{K/k,S,T}^V)$ belongs to* $\mathrm{im}(\nu)$, *and satisfies*

$$\mathcal{N}_H(\epsilon_{K/k,S,T}^V) = (-1)^{r(r'-r)} \cdot \nu(\mathrm{Rec}_W(\epsilon_{L/k,S,T}^{V'})).$$

Remark 5.5. In this article we write 'MRS($K/L/k, S, T$) is valid' to mean that the statement of Conjecture 5.4 is valid for all possible choices of V and V'.

Remark 5.6. In §6 we show that Conjecture 5.4 constitutes a natural refinement and generalization of both a conjecture of Darmon from [12] and of several conjectures of Gross from [21]. In addition, in a subsequent article [9] we will show that the validity of Conjecture 5.4 also implies the 'Gross-Stark conjecture' formulated by Gross in [20] and a refinement of the main result of Solomon in [47] concerning the 'wild Euler system' that he constructs in loc. cit.

Remark 5.7. One has $I(G_i)\mathbb{Z}[H] \subset I(H)$, so $J_W \subseteq I(H)^e$ where $e := r' - r \geq 0$. Thus there is a natural homomorphism

$$(\bigcap_G^r \mathcal{O}_{K,S,T}^\times) \otimes_\mathbb{Z} (J_W)_H \longrightarrow (\bigcap_G^r \mathcal{O}_{K,S,T}^\times) \otimes_\mathbb{Z} I(H)^e/I(H)^{e+1}.$$

Conjecture 5.4 is therefore a strengthening of the central conjecture of the third author in [46, Conj. 3] and of the conjecture formulated by Mazur and Rubin in [37, Conj. 5.2], both of which claim only that the given equality is valid after projection to the group $(\bigcap_G^r \mathcal{O}_{K,S,T}^\times) \otimes_\mathbb{Z} I(H)^e/I(H)^{e+1}$. This refinement is in the same spirit as Tate's strengthening in [51] of the 'refined class number formula' formulated by Gross in [21].

Remark 5.8. Note that, when $r = 0$, following [46, Def. 2.13] $\mathcal{N}_H$ would be defined to be the natural map $\mathbb{Z}[G] \to \mathbb{Z}[G]/I_H\mathcal{J}_W$, but this does not make any change because of the observation of Mazur and Rubin in [37, Lem. 5.6 (iv)]. Note also that, by Remark 4.13, the map $\mathbf{j}_{L/K}$ in [37, Lem. 4.9] (where our K/L is denoted by L/K) is essentially the same as our homomorphism ν. Finally we note that Mazur and Rubin do not use the fact that $\mathbf{j}_{L/K}$ is injective, so the formulation of [46, Conj. 3] is slightly stronger than the conjecture [37, Conj. 5.2].

We next state a refinement of a conjecture that was formulated by the first author in [3] (the original version of which has been studied in many subsequent articles of different authors including [22], [16], [17], [48], and [46]).

Conjecture 5.9. (B($K/L/k, S, T, V, V'$)). *For every*

$$\Phi \in \bigwedge_G^r \mathrm{Hom}_G(\mathcal{O}_{K,S,T}^\times, \mathbb{Z}[G]),$$

we have

$$\Phi(\epsilon_{K/k,S,T}^V) \in \mathcal{J}_W$$

and an equality

$$\Phi(\epsilon_{K/k,S,T}^V) = (-1)^{r(r'-r)}\Phi^H(\mathrm{Rec}_W(\epsilon_{L/k,S,T}^{V'})) \text{ in } (\mathcal{J}_W)_H.$$

In this article we improve an argument of the third author in [46] to prove the following result.

Theorem 5.10. *The conjectures*

$$\mathrm{MRS}(K/L/k, S, T, V, V') \text{ and } \mathrm{B}(K/L/k, S, T, V, V')$$

are equivalent.

The proof of this result will be given in §5.5.

5.4. An explicit resolution. As a preliminary step, we choose a useful representative of the complex

$$D^{\bullet}_{K,S,T} := R\Gamma_T((\mathcal{O}_{K,S})_{\mathcal{W}}, \mathbb{G}_m) \in D^{\mathrm{p}}(\mathbb{Z}[G]).$$

To do this we follow the method used in [3, §7].

Let d be a sufficiently large integer, and F be a free G-module of rank d with basis $b = \{b_i\}_{1\le i\le d}$. We define a surjection

$$\pi : F \longrightarrow \mathcal{S}^{\mathrm{tr}}_{S,T}(\mathbb{G}_{m/K})(= H^1(D^{\bullet}_{K,S,T}))$$

as follows. Recall that $S = \{v_0, \dots, v_n\}$. Let $F_{\le n}$ be a free $\mathbb{Z}[G]$-module generated by $\{b_i\}_{1\le i\le n}$. First, choose a homomorphism

$$\pi_1 : F_{\le n} \longrightarrow \mathcal{S}^{\mathrm{tr}}_{S,T}(\mathbb{G}_{m/K})$$

such that the composition map

$$F_{\le n} \xrightarrow{\pi_1} \mathcal{S}^{\mathrm{tr}}_{S,T}(\mathbb{G}_{m/K}) \longrightarrow X_{K,S}$$

sends b_i to $w_i - w_0$. (Such a homomorphism exists since $F_{\le n}$ is free.) Next, let A denote the kernel of the composition map

$$\mathcal{S}^{\mathrm{tr}}_{S,T}(\mathbb{G}_{m/K}) \longrightarrow X_{K,S} \longrightarrow Y_{K,S\setminus\{v_0\}},$$

where the last map sends the places above v_0 to 0. Since d is sufficiently large, we can choose a surjection

$$\pi_2 : F_{>n} \longrightarrow A,$$

where $F_{>n}$ is the free $\mathbb{Z}[G]$-module generated by $\{b_i\}_{n<i\le d}$. Define

$$\pi := \pi_1 \oplus \pi_2 : F = F_{\le n} \oplus F_{>n} \longrightarrow \mathcal{S}^{\mathrm{tr}}_{S,T}(\mathbb{G}_{m/K}).$$

One can easily show that π is surjective.

$D^{\bullet}_{K,S,T}$ defines a Yoneda extension class in $\mathrm{Ext}^2_G(\mathcal{S}^{\mathrm{tr}}_{S,T}(\mathbb{G}_{m/K}), \mathcal{O}^{\times}_{K,S,T})$. Since $D^{\bullet}_{K,S,T}$ is perfect, this class is represented by an exact sequence of the following form:

$$0 \longrightarrow \mathcal{O}^{\times}_{K,S,T} \longrightarrow P \xrightarrow{\psi} F \xrightarrow{\pi} \mathcal{S}^{\mathrm{tr}}_{S,T}(\mathbb{G}_{m/K}) \longrightarrow 0, \tag{24}$$

where π is the above map and P is a cohomologically-trivial G-module. Since $\mathcal{O}^{\times}_{K,S,T}$ is $\mathbb{Z}$-torsion-free, it follows that P is also $\mathbb{Z}$-torsion-free. Hence, P is projective. Note that the complex

$$P \xrightarrow{\psi} F,$$

where P is placed in degree 0, is quasi-isomorphic to $D^{\bullet}_{K,S,T}$. Hence we have an isomorphism

$$\det_G(D^{\bullet}_{K,S,T}) \simeq \det_G(P) \otimes_G \det_G^{-1}(F). \tag{25}$$

For each $1 \le i \le d$, we define

$$\psi_i := b_i^* \circ \psi \in \mathrm{Hom}_G(P, \mathbb{Z}[G]),$$

where $b_i^* \in \mathrm{Hom}_G(F, \mathbb{Z}[G])$ is the dual basis of $b_i \in F$.

5.5. The equivalence of Conjectures 5.4 and 5.9. In this subsection, we prove Theorem 5.10.

Proof of Theorem 5.10. We regard $\mathcal{O}^{\times}_{K,S,T} \subset P$ by the exact sequence (24). Note that, since $P/\mathcal{O}^{\times}_{K,S,T} \simeq \mathrm{im}(\psi) \subset F$ is $\mathbb{Z}$-torsion-free, we can apply Lemma 4.7 and Remark 4.8 for $M = \mathcal{O}^{\times}_{K,S,T}$. If $\mathcal{N}_H(\epsilon^V_{K/k,S,T}) \in \mathrm{im}(\nu)$, then we have

$$\Phi(\epsilon^V_{K/k,S,T}) = \Phi^H(\nu^{-1}(\mathcal{N}_H(\epsilon^V_{K/k,S,T}))) \text{ in } (\mathcal{J}_W)_H \tag{26}$$

for every $\Phi \in \bigwedge^r_G \mathrm{Hom}_G(\mathcal{O}^{\times}_{K,S,T}, \mathbb{Z}[G])$, by Proposition 4.17. Hence Conjecture 5.4 implies Conjecture 5.9.

Conversely, suppose that Conjecture 5.9 is valid. Then we have $\Phi(\epsilon^V_{K/k,S,T}) \in \mathcal{J}_W$ for every $\Phi \in \bigwedge^r_G \mathrm{Hom}_G(\mathcal{O}^{\times}_{K,S,T}, \mathbb{Z}[G])$, so again we use Proposition 4.17 to deduce that

$$\mathcal{N}_H(\epsilon^V_{K/k,S,T}) \in \mathrm{im}\left(\nu : (\bigwedge^r_{G/H} P^H) \otimes_{\mathbb{Z}} (J_W)_H \to (\bigwedge^r_G P) \otimes_{\mathbb{Z}} \mathbb{Z}[H]/I(H)J_W\right),$$

and that the equality (26) holds for every $\Phi \in \bigwedge^r_G \mathrm{Hom}_G(P, \mathbb{Z}[G])$. By Proposition 4.15, we see that the equality

$$\nu^{-1}(\mathcal{N}_H(\epsilon^V_{K/k,S,T})) = (-1)^{r(r'-r)}\mathrm{Rec}_W(\epsilon^{V'}_{L/k,S,T})$$

holds in $(\bigwedge^r_{G/H} P^H) \otimes_{\mathbb{Z}} (J_W)_H$. Since the natural map

$$(\bigcap^r_{G/H} \mathcal{O}^{\times}_{L,S,T}) \otimes_{\mathbb{Z}} (J_W)_H \longrightarrow (\bigwedge^r_{G/H} P^H) \otimes_{\mathbb{Z}} (J_W)_H$$

is injective by Remark 4.8, we see that the above equality holds in $(\bigcap^r_{G/H} \mathcal{O}^{\times}_{L,S,T}) \otimes_{\mathbb{Z}} (J_W)_H$. Thus Conjecture 5.9 implies Conjecture 5.4. $\square$

Remark 5.11. Although in the proof of Theorem 5.10 we used the exact sequence (24) to verify the existence of a finitely generated projective G-module P and an injection $\mathcal{O}^{\times}_{K,S,T} \hookrightarrow P$ whose cokernel is $\mathbb{Z}$-torsion-free, the referee pointed out that it is unnecessary to use (24) at this point. Indeed, choosing a projective module P' and a surjection $f : P' \to \mathrm{Hom}_{\mathbb{Z}}(\mathcal{O}^{\times}_{K,S,T}, \mathbb{Z})$, we have an embedding $\mathcal{O}^{\times}_{K,S,T} \hookrightarrow P := \mathrm{Hom}_{\mathbb{Z}}(P', \mathbb{Z})$ by dualizing f, whose cokernel is $\mathbb{Z}$-torsion-free.

5.6. The leading term conjecture implies the Rubin-Stark conjecture. The following result was first proved by the first author in [3, Cor. 4.1] but the proof given here is very much simpler than that given in loc. cit.

Theorem 5.12. LTC(K/k) *implies the Rubin-Stark conjecture for both* $(K/k,S,T,V)$ *and* $(L/k,S,T,V')$.

Proof. Assume that LTC(K/k) is valid so the zeta element $z_{K/k,S,T}$ is a $\mathbb{Z}[G]$-basis of $\det_G(D^\bullet_{K,S,T})$. In this case one also knows that P must be free of rank d and we define $z_b \in \bigwedge_G^d P$ to be the element corresponding to the zeta element $z_{K/k,S,T} \in \det_G(D^\bullet_{K,S,T})$ via the isomorphism

$$\bigwedge_G^d P \xrightarrow{\sim} \bigwedge_G^d P \otimes \bigwedge_G^d \mathrm{Hom}_G(F,\mathbb{Z}[G]) \simeq \det_G(D^\bullet_{K,S,T}),$$

where the first isomorphism is defined by

$$a \mapsto a \otimes \bigwedge_{1\le i\le d} b_i^*,$$

and the second isomorphism is given by (25).
Then Theorem 5.12 follows immediately from the next theorem (see also Corollary 5.15 below for $(L/k,S,T,V')$). □

Remark 5.13. In [52] Vallières closely follows the proof of [3, Cor. 4.1] to show that Conjecture 3.1 (and hence also LTC(K/k) by virtue of Proposition 3.4) implies the extension of the Rubin-Stark Conjecture formulated by Emmons and Popescu in [13]. The arguments used here can be used to show that LTC(K/k) implies a refinement of the main result of Vallières, and hence also of the conjecture of Emmons and Popescu, that is in the spirit of Theorem 1.5. This result is to be explained in forthcoming work of Livingstone-Boomla.

The following theorem was essentially obtained in [3] by the first author. This theorem describes the Rubin-Stark element in terms of the zeta elements. It is a key to prove Theorem 5.12, and also plays important roles in the proofs of Theorem 5.16 and Theorem 7.5 given below.

Theorem 5.14. *Assume that* LTC(K/k) *holds. Then, regarding* $\mathcal{O}^\times_{K,S,T}$ *as a submodule of* P*, one has*

$$(\bigwedge_{r<i\le d}\psi_i)(z_b) \in \bigcap_G^r \mathcal{O}^\times_{K,S,T} (\subset \bigwedge_G^r P)$$

(see Lemma 4.7 (ii)) and also

$$(-1)^{r(d-r)}(\bigwedge_{r<i\le d}\psi_i)(z_b) = \epsilon^V_{K/k,S,T}.$$

Proof. Take any $\chi \in \widehat{G}$. Recall from (21) that

$$r_{\chi,S} = \dim_{\mathbb{C}}(e_\chi \mathbb{C}X_{K,S}) = \dim_{\mathbb{C}}(e_\chi \mathbb{C}\mathcal{O}^\times_{K,S,T})$$

(the last equality follows from $\mathbb{C}\mathcal{O}_{K,S,T}^{\times} \simeq \mathbb{C}X_{K,S}$). Consider the map

$$\Psi := \bigoplus_{r<i\le d} \psi_i : e_\chi \mathbb{C}P \longrightarrow e_\chi \mathbb{C}[G]^{\oplus(d-r)}.$$

This map is surjective if and only if $r_{\chi,S} = r$. Indeed, if $r_{\chi,S} = r$, then $\{e_\chi(w_i - w_0)\}_{1\le i\le r}$ is a $\mathbb{C}$-basis of $e_\chi \mathbb{C}X_{K,S}$, so $e_\chi \mathbb{C}\,\mathrm{im}(\psi) = e_\chi \mathbb{C}\ker(\pi) = \bigoplus_{r<i\le d} e_\chi \mathbb{C}[G]b_i$. In this case, Ψ is surjective. If $r_{\chi,S} > r$, then $\dim_{\mathbb{C}}(e_\chi \mathbb{C}\,\mathrm{im}(\psi)) = d - r_{\chi,S} < d - r$, so Ψ is not surjective. Applying Lemma 4.2, we have

$$e_\chi(\bigwedge_{r<i\le d} \psi_i)(z_b) \begin{cases} \in e_\chi \mathbb{C} \bigwedge_G^r \mathcal{O}_{K,S,T}^{\times}, & \text{if } r_{\chi,S} = r, \\ = 0, & \text{if } r_{\chi,S} > r. \end{cases}$$

From this and Lemma 4.7 (ii), we have

$$(\bigwedge_{r<i\le d} \psi_i)(z_b) \in (\mathbb{Q}\bigwedge_G^r \mathcal{O}_{K,S,T}^{\times}) \cap \bigwedge_G^r P = \bigcap_G^r \mathcal{O}_{K,S,T}^{\times}.$$

By Lemma 4.3 and the definition of z_b, we have

$$\lambda_{K,S}((-1)^{r(d-r)}(\bigwedge_{r<i\le d} \psi_i)(z_b)) = \theta_{K/k,S,T}^{(r)} \bigwedge_{1\le i\le r} (w_i - w_0).$$

By the characterization of the Rubin-Stark element, we have

$$(-1)^{r(d-r)}(\bigwedge_{r<i\le d} \psi_i)(z_b) = \epsilon_{K/k,S,T}^{V}.$$

This completes the proof. □

By the same argument as above, one obtains the following result.

Corollary 5.15. *Assume that* LTC(K/k) *holds. Then we have an equality*

$$(-1)^{r'(d-r')}(\bigwedge_{r'<i\le d} \psi_i^H)(\mathrm{N}_H^d\, z_b) = \epsilon_{L/k,S,T}^{V'}$$

in $\bigcap_{G/H}^{r'} \mathcal{O}_{L,S,T}^{\times}$.

5.7. The leading term conjecture implies Conjecture 5.4. In this subsection we prove the following result.

Theorem 5.16. LTC(K/k) *implies* MRS$(K/L/k,S,T,V,V')$.

By Remark 3.3, this directly implies the following result.

Corollary 5.17. MRS$(K/L/k,S,T,V,V')$ *is valid if K is an abelian extension over $\mathbb{Q}$ or if k is a function field.*

Remark 5.18. Theorem 5.16 is an improvement of the main result in [46, Th. 3.22] by the third author, which asserts that under some hypotheses LTC(K/k) implies most of Conjecture 5.4. In [3, Th. 3.1], the first author proved that LTC(K/k) implies most of Conjecture 5.9. Since we know by Theorem 5.10

that Conjecture 5.4 and Conjecture 5.9 are equivalent, Theorem 5.16 is also an improvement of [3, Th. 3.1].

Remark 5.19. In [46, §4], by using a weak version of Corollary 5.17, the third author gave another proof of the 'except 2-part' of Darmon's conjecture on cyclotomic units [12], which was first proved by Mazur and Rubin in [36] via Kolyvagin systems. In §6, we shall use Corollary 5.17 to give a full proof of a refined version of Darmon's conjecture, and also give a new evidence for Gross's conjecture on tori [21], which was studied by Hayward [22], Greither and Kučera [16], [17].

We prove Theorem 5.16 after proving some lemmas. The following lemma is a restatement of [3, Lem. 7.4].

Lemma 5.20. *If $1 \leq i \leq n$, then we have an inclusion*

$$\mathrm{im}(\psi_i) \subset \mathcal{I}_i.$$

In particular, $\psi_i = 0$ for $1 \leq i \leq r$.

Proof. Take any $a \in P$. Write

$$\psi(a) = \sum_{j=1}^{d} x_j b_j$$

with some $x_j \in \mathbb{Z}[G]$. For each i with $1 \leq i \leq n$, we show that $x_i \in \mathcal{I}_i$, or equivalently, $\mathrm{N}_{G_i} x_i = 0$. Noting that F^{G_i} is a free G/G_i-module with basis $\{\mathrm{N}_{G_i} b_j\}_{1 \leq j \leq d}$, it is sufficient to show that

$$\sum_{j=1}^{d} \mathrm{N}_{G_i} x_j b_j \in \langle \mathrm{N}_{G_i} b_j : 1 \leq j \leq d, j \neq i \rangle_{G/G_i}.$$

The left hand side is equal to $\psi(\mathrm{N}_{G_i} a)$. By the exact sequence (24), this is contained in $\ker(\pi|_{F^{G_i}})$. Note that we have a natural isomorphism

$$\mathrm{N}_{G_i} X_{K,S} \simeq X_{K^{G_i},S}.$$

Since v_i splits completely in K^{G_i}, the G/G_i-submodule of $\mathrm{N}_{G_i} X_{K,S}$ generated by $\mathrm{N}_{G_i}(w_i - w_0)$ is isomorphic to $\mathbb{Z}[G/G_i]$. This shows that

$$\ker(\pi|_{F^{G_i}}) \subset \langle \mathrm{N}_{G_i} b_j : 1 \leq j \leq d, j \neq i \rangle_{G/G_i}.$$

□

For each integer i with $r < i \leq r'$, we define a map

$$\widetilde{\mathrm{Rec}}_i : P^H \longrightarrow (\mathcal{I}_i)_H$$

as follows. For $a \in P^H$, take $\widetilde{a} \in P$ such that $\mathrm{N}_H \widetilde{a} = a$ (this is possible since P is cohomologically-trivial). Define

$$\widetilde{\mathrm{Rec}}_i(a) := \psi_i(\widetilde{a}) \bmod I_H \mathcal{I}_i \ \in (\mathcal{I}_i)_H.$$

(Note that $\mathrm{im}(\psi_i) \subset \mathcal{I}_i$ by Lemma 5.20.) One can easily check that this is well-defined.

Lemma 5.21. *On $\mathcal{O}_{L,S,T}^{\times}$, which we regard as a submodule of P^H, $\widetilde{\mathrm{Rec}}_i$ coincides with the map Rec_i. In particular, by the construction of (20), we can extend the map*

$$\mathrm{Rec}_W : \bigcap_{G/H}^{r'} \mathcal{O}_{L,S,T}^{\times} \longrightarrow (\bigcap_{G/H}^{r} \mathcal{O}_{L,S,T}^{\times}) \otimes_{\mathbb{Z}} (J_W)_H$$

to

$$\widetilde{\mathrm{Rec}}_W := \bigwedge_{r<i\leq r'} \widetilde{\mathrm{Rec}}_i : \bigwedge_{G/H}^{r'} P^H \to (\bigwedge_{G/H}^{r} P^H)\otimes_{G/H}(\mathcal{J}_W)_H \simeq (\bigwedge_{G/H}^{r} P^H)\otimes_{\mathbb{Z}}(J_W)_H\ .$$

Proof. The proof is essentially the same as [3, Prop. 10.1] and [2, Lem. 8]. For $a \in \mathcal{O}_{L,S,T}^{\times}$, take $\widetilde{a} \in P$ such that $\mathrm{N}_H\,\widetilde{a} = a$ in P^H. For each $\tau \in G/H$, fix a lift $\widetilde{\tau} \in G$. Regard F as the free H-module with basis $\{\widetilde{\tau} b_i\}_{i,\tau}$. It is sufficient to show that

$$(\widetilde{\tau} b_i)^* \circ \psi(\widetilde{a}) = \mathrm{rec}_{\widetilde{\tau} w_i}(a) - 1 = \mathrm{rec}_{w_i}(\tau^{-1}a) - 1, \tag{27}$$

for every $r < i \leq r'$, where $(\widetilde{\tau} b_i)^* \in \mathrm{Hom}_H(F, \mathbb{Z}[H])$ is the dual basis of F as a free H-module. Indeed, using

$$\widetilde{\mathrm{Rec}}_i(a) = \psi_i(\widetilde{a}) = \sum_{\tau \in G/H} \widetilde{\tau}((\widetilde{\tau} b_i)^* \circ \psi(\widetilde{a})),$$

we know from (27) that

$$\widetilde{\mathrm{Rec}}_i(a) = \sum_{\tau \in G/H} \widetilde{\tau}(\mathrm{rec}_{w_i}(\tau^{-1}a) - 1) = \mathrm{Rec}_i(a).$$

We shall show (27). For simplicity, set $w := \widetilde{\tau} w_i$ and $b := \widetilde{\tau} b_i$. We denote the decomposition group of w by $G_w (= G_i)$. As in the proof of Proposition 2.4, one can show that there is a unique morphism

$$\theta_w : \mathbb{Q}[-2] \longrightarrow R\Gamma(K_w, \mathbb{G}_m)$$

in $D(\mathbb{Z}[G_w])$ such that $H^2(\theta_w)$ is equal to the natural map

$$\mathbb{Q} \longrightarrow \mathbb{Q}/\mathbb{Z} \simeq H^2(K_w, \mathbb{G}_m),$$

where the last isomorphism is the invariant map in the local class field theory. We define the complex $R\Gamma((K_w)_{\mathcal{W}}, \mathbb{G}_m)$ by

$$R\Gamma((K_w)_{\mathcal{W}}, \mathbb{G}_m) := \mathrm{Cone}(\theta_w)$$

for the local field K_w. We have natural identifications $H^0((K_w)_{\mathcal{W}}, \mathbb{G}_m) = K_w^{\times}$ and $H^1((K_w)_{\mathcal{W}}, \mathbb{G}_m) = \mathbb{Z}$. The complex $R\Gamma((K_w)_{\mathcal{W}}, \mathbb{G}_m)$ defines a Yoneda extension class τ_w in $\mathrm{Ext}^2_{G_w}(\mathbb{Z}, K_w^{\times})$, and [6, Prop. 3.5(a)] shows that τ_w coincides with the local fundamental class in $H^2(G_w, K_w^{\times})$. The class τ_w is represented by a 2-extension of the form

$$0 \longrightarrow K_w^{\times} \longrightarrow P_w \xrightarrow{\psi_w} \mathbb{Z}[G_w] \longrightarrow \mathbb{Z} \longrightarrow 0,$$

where P_w is a cohomologically trivial G_w-module. Define

$$\rho_w : L_w^\times \longrightarrow I(G_w)/I(G_w)^2$$

by $\rho(x) := \psi_w(\widetilde{x})$, where $\widetilde{x} \in P_w$ is taken so that $\mathrm{N}_{G_w}\, \widetilde{x} = x$ (note that $\psi_w(\widetilde{x})$ is well-defined in $I(G_w)/I(G_w)^2$). Then, the existence of a natural localization morphism

$$D^\bullet_{K,S,T} = R\Gamma_T((\mathcal{O}_{K,S})_\mathcal{W}, \mathbb{G}_m) \longrightarrow \mathbb{Z}[H] \otimes^{\mathbb{L}}_{\mathbb{Z}[G_w]} R\Gamma((K_w)_\mathcal{W}, \mathbb{G}_m)$$

and our choice of a representative of $D^\bullet_{K,S,T}$ implies

$$b^* \circ \psi(\widetilde{a}) = \rho_w(a) \text{ in } I(G_w)/I(H)I(G_w).$$

Hence, (27) is reduced to the equality

$$\rho_w = \mathrm{rec}_w - 1.$$

Consider the map

$$I(G_w)/I(G_w)^2 = \widehat{H}^{-1}(G_w, I(G_w)) \simeq \widehat{H}^{-2}(G_w, \mathbb{Z}) \xrightarrow{\sim} \widehat{H}^0(G_w, K_w^\times), \tag{28}$$

where the first isomorphism is the connecting homomorphism with respect to the short exact sequence

$$0 \longrightarrow I(G_w) \longrightarrow \mathbb{Z}[G_w] \longrightarrow \mathbb{Z} \longrightarrow 0,$$

and the last is given by the cup product with τ_w. The map (28) is the inverse of $\mathrm{rec}_w - 1$ by definition. One can also check that (28) coincides with the δ-map of the snake lemma applied to the diagram

$$\begin{tikzcd}
 & (K_w^\times)_{G_w} \arrow[r] \arrow[d, "\mathrm{N}_{G_w}"] & (P_w)_{G_w} \arrow[r, "\psi_w"] \arrow[d, "\mathrm{N}_{G_w}"] & I(G_w)/I(G_w)^2 \arrow[r] \arrow[d, "0"] & 0 \\
0 \arrow[r] & L_w^\times \arrow[r] & P_w^{G_w} \arrow[r, "\psi_w"] & I(G_w)^{G_w},
\end{tikzcd}$$

i.e. the inverse of ρ_w. Thus we have $\rho_w = \mathrm{rec}_w - 1$, which completes the proof. □

Note that, by Lemma 5.20, $\bigwedge_{r<i\leq d} \psi_i$ defines a map

$$\bigwedge_G^d P \longrightarrow \mathcal{J}_W \bigwedge_G^r P.$$

Let ν be the injection

$$\nu : (\bigwedge_{G/H}^r P^H) \otimes_\mathbb{Z} (J_W)_H \longrightarrow (\bigwedge_G^r P) \otimes_\mathbb{Z} \mathbb{Z}[H]/I(H)J_W$$

in Proposition 4.12. By Proposition 4.17, we have

$$\mathcal{N}_H(\mathcal{J}_W \bigwedge_G^r P) \subset \mathrm{im}(\nu),$$

so we can define a map

$$\nu^{-1} \circ \mathcal{N}_H : \mathcal{J}_W \bigwedge_G^r P \longrightarrow (\bigwedge_{G/H}^r P^H) \otimes_{\mathbb{Z}} (J_W)_H.$$

Lemma 5.22. *We have the following commutative diagram:*

$$\begin{tikzcd}
\bigwedge_G^d P \arrow[r] \arrow[d, "\mathrm{N}_H^d"'] & \mathcal{J}_W \bigwedge_G^r P \arrow[d, "\nu^{-1}\circ\mathcal{N}_H"] \\
\bigwedge_{G/H}^d P^H \arrow[r] & (\bigwedge_{G/H}^r P^H) \otimes_{\mathbb{Z}} (J_W)_H,
\end{tikzcd}$$

where the top arrow is $(-1)^{r(d-r)} \bigwedge_{r<i\le d} \psi_i$*, and the bottom arrow is the composition of* $(-1)^{r(r'-r)}\widetilde{\mathrm{Rec}}_W$ *and* $(-1)^{r'(d-r')} \bigwedge_{r'<i\le d} \psi_i^H$.

Proof. We can prove this lemma by explicit computations, using Proposition 4.1, Proposition 4.11, and Remark 4.14. $\square$

Proof of Theorem 5.16. By Remark 4.8 we may compute in $(\bigwedge_{G/H}^r P^H) \otimes_{\mathbb{Z}} (J_W)_H$. Using Corollary 5.15, Lemma 5.21, Lemma 5.22, and Theorem 5.14 in this order, we compute

$$\begin{aligned}
& (-1)^{r(r'-r)}\mathrm{Rec}_W(\epsilon^{V'}_{L/k,S,T}) \\
= \;& (-1)^{r(r'-r)}\widetilde{\mathrm{Rec}}_W((-1)^{r'(d-r')}(\bigwedge_{r'<i\le d} \psi_i^H)(\mathrm{N}_H^d\, z_b)) \\
= \;& (-1)^{r(d-r)}\nu^{-1}(\mathcal{N}_H((\bigwedge_{r<i\le d} \psi_i)(z_b))) \\
= \;& \nu^{-1}(\mathcal{N}_H(\epsilon^V_{K/k,S,T})).
\end{aligned}$$

This completes the proof of Theorem 5.16. $\square$

6. Conjectures of Darmon and of Gross

In this section we use Corollary 5.17 to prove a refined version of the conjecture formulated by Darmon in [12] and to obtain important new evidence for a refined version of the 'conjecture for tori' formulated by Gross in [21].

6.1. Darmon's Conjecture. We formulate a slightly modified and refined version of Damon's conjecture ([12],[36]).

Let L be a real quadratic field. Let f be the conductor of L. Let χ be the Dirichlet character defined by

$$\chi : (\mathbb{Z}/f\mathbb{Z})^\times = \mathrm{Gal}(\mathbb{Q}(\mu_f)/\mathbb{Q}) \longrightarrow \mathrm{Gal}(L/\mathbb{Q}) \simeq \{\pm 1\},$$

where the first map is the restriction map. Fix a square-free positive integer n which is coprime to f, and let K be the maximal real subfield of $L(\mu_n)$. Set $G := \mathrm{Gal}(K/\mathbb{Q})$ and $H := \mathrm{Gal}(K/L)$. Put $n_\pm := \prod_{\ell | n, \chi(\ell)=\pm 1} \ell$, and

$\nu_\pm := |\{\ell | n_\pm\}|$ (in this section, ℓ always denotes a prime number). We fix an embedding $\overline{\mathbb{Q}} \hookrightarrow \mathbb{C}$. Define a cyclotomic unit by

$$\beta_n := \mathrm{N}_{L(\mu_n)/K}(\prod_{\sigma \in \mathrm{Gal}(\mathbb{Q}(\mu_{nf})/\mathbb{Q}(\mu_n))} \sigma(1-\zeta_{nf})^{\chi(\sigma)}) \in K^\times,$$

where $\zeta_{nf} = e^{\frac{2\pi i}{nf}}$. Let τ be the generator of $G/H = \mathrm{Gal}(L/\mathbb{Q})$. Write $n_+ = \ell_1 \cdots \ell_{\nu_+}$. Note that $(1-\tau)\mathcal{O}_L[1/n]^\times$ is a free abelian group of rank $\nu_+ + 1$ (see [36, Lem. 3.2 (ii)]). Take $u_0, \ldots, u_{\nu_+} \in \mathcal{O}_L[1/n]^\times$ so that $\{u_0^{1-\tau}, \ldots, u_{\nu_+}^{1-\tau}\}$ is a basis of $(1-\tau)\mathcal{O}_L[1/n]^\times$ and that $\det(\log |u_i^{1-\tau}|_{\lambda_j})_{0 \le i,j \le \nu_+} > 0$, where each λ_j $(1 \le j \le \nu_+)$ is a (fixed) place of L lying above ℓ_j, and λ_0 is the infinite place of L determined by the embedding $\overline{\mathbb{Q}} \hookrightarrow \mathbb{C}$ fixed above. Define

$$R_n := (\bigwedge_{1 \le i \le \nu_+} (\mathrm{rec}_{\lambda_i} - 1))(u_0^{1-\tau} \wedge \cdots \wedge u_{\nu_+}^{1-\tau}) \in L^\times \otimes_{\mathbb{Z}} (J_{n_+})_H,$$

where

$$J_{n_+} := \begin{cases} (\prod_{i=1}^{i=\nu_+} I(G_{\ell_i}), & \text{if } \nu_+ \neq 0, \\ \mathbb{Z}[H], & \text{if } \nu_+ = 0, \end{cases}$$

where G_{ℓ_i} is the decomposition group of ℓ_i in G (note that since ℓ_i splits in L, we have $G_{\ell_i} \subset H$), and $(J_{n_+})_H := J_{n_+}/I(H)J_{n_+}$. We set $h_n := |\mathrm{Pic}(\mathcal{O}_L[\frac{1}{n}])|$. For any element $a \in K^\times$, following Definition 4.16 we define

$$\mathcal{N}_H(a) := \sum_{\sigma \in H} \sigma a \otimes \sigma^{-1} \in K^\times \otimes_{\mathbb{Z}} \mathbb{Z}[H]/I(H)J_{n_+}.$$

Note that, since $K^\times/L^\times$ is $\mathbb{Z}$-torsion-free, the natural map

$$(L^\times/\{\pm 1\}) \otimes_{\mathbb{Z}} (J_{n_+})_H \longrightarrow (K^\times/\{\pm 1\}) \otimes_{\mathbb{Z}} \mathbb{Z}[H]I(H)J_{n_+}$$

is injective.

Our refined Darmon's conjecture is formulated as follows.

Theorem 6.1. *One has*

$$\mathcal{N}_H(\beta_n) = -2^{\nu_-} h_n R_n \quad \textit{in} \quad (L^\times/\{\pm 1\}) \otimes_{\mathbb{Z}} (J_{n_+})_H.$$

Remark 6.2. Let I_n be the augmentation ideal of $\mathbb{Z}[\mathrm{Gal}(L(\mu_n)/L)]$. Note that there is a natural isomorphism

$$I_n^{\nu_+}/I_n^{\nu_+ + 1} \otimes_{\mathbb{Z}} \mathbb{Z}[\frac{1}{2}] \xrightarrow{\sim} I(H)^{\nu_+}/I(H)^{\nu_+ + 1} \otimes_{\mathbb{Z}} \mathbb{Z}[\frac{1}{2}].$$

It is not difficult to see that the following statement is equivalent to [36, Th. 3.9]:

$$\mathcal{N}_H(\beta_n) = -2^{\nu_-} h_n R_n \quad \textit{in} \quad (L^\times/\{\pm 1\}) \otimes_{\mathbb{Z}} I(H)^{\nu_+}/I(H)^{\nu_+ + 1} \otimes_{\mathbb{Z}} \mathbb{Z}[\frac{1}{2}]$$

(see [46, Lem. 4.7]). Since there is a natural map $(J_{n_+})_H \longrightarrow I(H)^{\nu_+}/I(H)^{\nu_+ + 1}$, Theorem 6.1 refines [36, Th. 3.9]. Note also that, in the original statement of Darmon's conjecture, the cyclotomic unit is defined by

$$\alpha_n := \prod_{\sigma \in \mathrm{Gal}(\mathbb{Q}(\mu_{nf})/\mathbb{Q}(\mu_n))} \sigma(1-\zeta_{nf})^{\chi(\sigma)},$$

whereas our cyclotomic unit is $\beta_n = \mathrm{N}_{L(\mu_n)/K}(\alpha_n)$. Since cyclotomic units, as Stark elements, lie in real fields, it is natural to consider β_n. Thus, modifying the original statement of Darmon's conjecture in the '2-part', we obtained Theorem 6.1, which does not exclude the '2-part'.

Proof of Theorem 6.1. We show that Darmon's conjecture is a consequence of Conjecture 5.4, and use Corollary 5.17 to prove it. We fit notation in this section into that in §5. Set $S := \{\infty\} \cup \{\ell | nf\}$. Take a prime v_0 of $\mathbb{Q}$, which divides f. We denote by w_1 the infinite place of K (and also L) which corresponds to the fixed embedding $\mathbb{Q} \hookrightarrow \mathbb{C}$. For $2 \le i \le \nu_+ + 1$, set $w_i := \lambda_{i-1}$. Let T be a finite set of primes that is disjoint from S and satisfying that $\mathcal{O}_{K,S,T}^\times$ is $\mathbb{Z}$-torsion-free. (In the sequel, we refer such a set of primes as 'T'.) Since K and L are abelian over $\mathbb{Q}$, the Rubin-Stark conjecture for $K/\mathbb{Q}$ and $L/\mathbb{Q}$ holds (see Remark 5.3 (iii)). Set $V := \{\infty\}$ and $V' := \{\infty, \ell_1, \ldots, \ell_{\nu_+}\}$. We denote $\epsilon_{K,T} = \epsilon^V_{K/\mathbb{Q},S,T} \in \mathcal{O}^\times_{K,S,T}$ and $\epsilon_{L,T} = \epsilon^{V'}_{L/\mathbb{Q},S,T} \in \bigcap_{G/H}^{\nu_+ + 1} \mathcal{O}^\times_{L,S,T}$ for the Rubin-Stark elements, characterized by

$$\lambda_{K,S}(\epsilon_{K,T}) = \theta^{(1)}_{K/\mathbb{Q},S,T}(w_1 - w_0),$$

$$\lambda_{L,S}(\epsilon_{L,T}) = \theta^{(\nu_+ + 1)}_{L/\mathbb{Q},S,T} \bigwedge_{1 \le i \le \nu_+ + 1} (w_i - w_0).$$

We take $\mathscr{T}$, a finite family of 'T', such that

$$\sum_{T \in \mathscr{T}} a_T \delta_T = 2$$

for some $a_T \in \mathbb{Z}[G]$, where $\delta_T := \prod_{\ell \in T}(1 - \ell \mathrm{Fr}_\ell^{-1})$ (see [50, Chap. IV, Lem. 1.1]). By [46, Lem. 4.6], we have

$$(1-\tau) \sum_{T \in \mathscr{T}} a_T \epsilon_{K,T} = \beta_n \quad in \quad K^\times/\{\pm 1\},$$

(where $\tau \in \mathrm{Gal}(L/\mathbb{Q})$ is regarded as an element of $\mathrm{Gal}(K/\mathbb{Q}(\mu_n)^+)$) and

$$(1-\tau) \sum_{T \in \mathscr{T}} a_T \epsilon_{L,T} = (-1)^{\nu_+ + 1} 2^{\nu_-} h_n (1-\tau) u_0 \wedge \cdots \wedge u_{\nu_+} \quad in \quad \mathbb{Q} \bigwedge_{G/H}^{\nu_+ + 1} \mathcal{O}^\times_{L,S}.$$

As in §5.3, for $1 < i \le \nu_+ + 1$ we denote by Rec_i the homomorphism

$$\mathrm{Rec}_i : \mathcal{O}^\times_{L,S,T} \longrightarrow (\mathcal{J}_{n_+})_H = J_{n_+}\mathbb{Z}[G]/I_H J_{n_+}\mathbb{Z}[G]$$

defined by

$$\mathrm{Rec}_i(a) = \mathrm{rec}_{\lambda_{i-1}}(a) - 1 + \tau(\mathrm{rec}_{\lambda_{i-1}}(\tau a) - 1).$$

$\bigwedge_{1 < i \le \nu_+ + 1} \mathrm{Rec}_i$ induces a homomorphism

$$\bigcap_{G/H}^{\nu_+ + 1} \mathcal{O}^\times_{L,S,T} \longrightarrow (\bigcap_{G/H}^{1} \mathcal{O}^\times_{L,S,T}) \otimes_{\mathbb{Z}} (J_{n_+})_H = \mathcal{O}^\times_{L,S,T} \otimes_{\mathbb{Z}} (J_{n_+})_H,$$

which we denote by Rec_{n_+}. We compute

$$\begin{aligned}
& (1-\tau)\sum_{T\in\mathscr{T}} a_T \mathrm{Rec}_{n_+}(\epsilon_{L,T}) \\
= & \sum_{T\in\mathscr{T}} \mathrm{Rec}_{n_+}(a_T(1-\tau)\epsilon_{L,T}) \\
= & \sum_{T\in\mathscr{T}} (\bigwedge_{1\le i\le \nu_+} (\mathrm{rec}_{\lambda_i}-1))((1-\tau)^{\nu_+ +1} a_T \epsilon_{L,T}) \\
= & (\bigwedge_{1\le i\le \nu_+} (\mathrm{rec}_{\lambda_i}-1))((-1)^{\nu_+ +1} 2^{\nu_-} h_n u_0^{1-\tau}\wedge\cdots\wedge u_{\nu_+}^{1-\tau}) \\
= & (-1)^{\nu_+ +1} 2^{\nu_-} h_n R_n.
\end{aligned}$$

By Corollary 5.17, we have

$$\mathcal{N}_H(\epsilon_{K,T}) = (-1)^{\nu_+}\mathrm{Rec}_{n_+}(\epsilon_{L,T})$$

(note that the map ν in Conjecture 5.4 is the natural inclusion map in this case.) Hence, we have

$$\begin{aligned}
\mathcal{N}_H(\beta_n) &= (1-\tau)\sum_{T\in\mathscr{T}} a_T \mathcal{N}_H(\epsilon_{K,T}) \\
&= (-1)^{\nu_+}(1-\tau)\sum_{T\in\mathscr{T}} a_T \mathrm{Rec}_{n_+}(\epsilon_{L,T}) \\
&= -2^{\nu_-} h_n R_n,
\end{aligned}$$

as required. □

6.2. Gross's conjecture for tori. In this section we use Corollary 5.17 to obtain some new evidence in support of the 'conjecture for tori' formulated by Gross in [21].

We review the formulation of Gross's conjecture for tori. We follow [22, Conj. 7.4]. Let k be a global field, and L/k be a quadratic extension. Let $\widetilde{L}/k$ be a finite abelian extension, which is disjoint to L, and set $K := L\widetilde{L}$. Set $G := \mathrm{Gal}(K/k)$, and $H := \mathrm{Gal}(K/L) = \mathrm{Gal}(\widetilde{L}/k)$. Let τ be the generator of $G/H = \mathrm{Gal}(L/k)$. Let S be a non-empty finite set of places of k such that $S_\infty(k)\cup S_{\mathrm{ram}}(K/k)\subset S$. Let T be a finite set of places of k that is disjoint from S and satisfies that $\mathcal{O}_{K,S,T}^\times$ is $\mathbb{Z}$-torsion-free. Let $v_1,\ldots,v_{r'}$ be all places in S which split in L. We assume $r' < |S|$. Then, by [45, Lem. 3.4 (i)], we see that $h_{k,S,T} := |\mathrm{Cl}_S^T(k)|$ divides $h_{L,S,T} := |\mathrm{Cl}_S^T(L)|$. Take $u_1,\ldots,u_{r'}\in\mathcal{O}_{L,S,T}^\times$ such that $\{u_1^{1-\tau},\ldots,u_{r'}^{1-\tau}\}$ is a basis of $(1-\tau)\mathcal{O}_{L,S,T}^\times$, which is isomorphic to $\mathbb{Z}^{\oplus r'}$, and $\det(-\log|u_i^{1-\tau}|_{w_j})_{1\le i,j\le r'} > 0$, where w_j is a (fixed) place of L lying above v_j. Put $W := \{v_1,\ldots,v_{r'}\}$. As in §5.3, we define

$$J_W := \begin{cases} (\prod_{0<i\le r'} I(G_i))\mathbb{Z}[H], & \text{if } W\neq\emptyset, \\ \mathbb{Z}[H], & \text{if } W=\emptyset, \end{cases}$$

where $G_i \subset H$ denotes the decomposition group of v_i, and $I(G_i)$ is the augmentation ideal of $\mathbb{Z}[G_i]$. Set

$$R_{S,T} := \det(\mathrm{rec}_{w_j}(u_i^{1-\tau}) - 1)_{1\le i,j\le r'} \in (J_W)_H.$$

Let χ be the non-trivial character of G/H. The map

$$\mathbb{Z}[G] = \mathbb{Z}[H \times G/H] \longrightarrow \mathbb{Z}[H]$$

induced by χ is also denoted by χ.
Gross's tori conjecture is formulated as follows.

Conjecture 6.3.

$$\chi(\theta_{K/k,S,T}(0)) = 2^{|S|-1-r'} \frac{h_{L,S,T}}{h_{k,S,T}} R_{S,T} \quad \mathit{in} \quad (J_W)_H.$$

Remark 6.4. The statement that the equality of Conjecture 6.3 holds in $\mathbb{Z}[H]/I(H)^{r'+1}$ is equivalent to [22, Conj. 7.4] (if we neglect the sign). Indeed, we see that

$$R_{S,T} = ((\mathcal{O}_{L,S,T}^\times)^- : (1-\tau)\mathcal{O}_{L,S,T}^\times)R_H^-,$$

where R_H^- is the 'minus-unit regulator' defined in [22, §7.2] (where our H is denoted by G). Since there is a natural map $(J_W)_H \to \mathbb{Z}[H]/I(H)^{r'+1}$, Conjecture 6.3 refines [22, Conj. 7.4].

Theorem 6.5. *Conjecture 5.4 implies Conjecture 6.3. In particular, Conjecture 6.3 is valid if K is an abelian extension over $\mathbb{Q}$ or k is a global function field.*

Proof. First, note that the Rubin-Stark conjecture for $(K/k, S, T, \emptyset)$ and $(L/k, S, T, W)$ is true by Remark 5.3 (i) and (ii), respectively. By Conjecture 5.4, we have

$$\theta_{K/k,S,T}(0) = \mathrm{Rec}_W(\epsilon_{L/k,S,T}^W) \quad \mathit{in} \quad (\mathcal{J}_W)_H (= \mathbb{Z}[G/H] \otimes_{\mathbb{Z}} (J_W)_H).$$

(Note that $\nu^{-1}(\mathcal{N}_H(\theta_{K/k,S,T}(0))) = \theta_{K/k,S,T}(0)$ in $(\mathcal{J}_W)_H$ by [37, Lem. 5.6 (iv)].) Note that $\chi \circ \mathrm{Rec}_i = \mathrm{rec}_{w_i}((1-\tau)(\cdot)) - 1$. So we have

$$\chi(\mathrm{Rec}_W(\epsilon_{L/k,S,T}^W)) = (\bigwedge_{1\le i\le r'} (\mathrm{rec}_{w_i} - 1))((1-\tau)^{r'} \epsilon_{L/k,S,T}^W).$$

We know by the proof of [45, Th. 3.5] that

$$(1-\tau)^{r'} \epsilon_{L/k,S,T}^W = 2^{|S|-1-r'} \frac{h_{L,S,T}}{h_{k,S,T}} u_1^{1-\tau} \wedge \cdots \wedge u_{r'}^{1-\tau}.$$

Hence we have

$$\begin{aligned}
\chi(\theta_{K/k,S,T}(0)) &= \chi(\mathrm{Rec}_W(\epsilon^W_{L/k,S,T}))\\
&= (\bigwedge_{1\le i\le r'}(\mathrm{rec}_{w_i}-1))((1-\tau)^{r'}\epsilon^W_{L/k,S,T})\\
&= 2^{|S|-1-r'}\frac{h_{L,S,T}}{h_{k,S,T}}(\bigwedge_{1\le i\le r'}(\mathrm{rec}_{w_i}-1))(u_1^{1-\tau}\wedge\cdots\wedge u_{r'}^{1-\tau})\\
&= 2^{|S|-1-r'}\frac{h_{L,S,T}}{h_{k,S,T}}R_{S,T},
\end{aligned}$$

as required.

Having now proved the first claim, the second claim follows directly from Corollary 5.17. □

Remark 6.6. The strongest previous evidence in favour of the conjecture for tori is that obtained by Greither and Kučera in [16, 17], in which it is referred to as the 'Minus Conjecture' and studied in a slightly weaker form in order to remove any occurence of the auxiliary set T. More precisely, by using rather different methods they were able to prove that this conjecture was valid in the case that $k=\mathbb{Q}$, $K=FK^+$ where F is imaginary quadratic of conductor f and class number h_F and $K^+/\mathbb{Q}$ is tamely ramified, abelian of exponent equal to an odd prime ℓ and ramified at precisely s primes $\{p_i\}_{1\le i\le s}$ each of which splits in $F/\mathbb{Q}$; further, any of the following conditions are satisfied

- $s=1$ and $\ell\nmid f$ [16, Th. 8.8], or
- $s=2, \ell\nmid fh_F$ and either $K^+/\mathbb{Q}$ is cyclic or p_1 is congruent to an ℓ-th power modulo p_2 [16, Th. 8.9], or
- $\ell\ge 3(s+1)$ and $\ell\nmid h_F$ [17, Th. 3.7].

It is straightforward to show that the conjecture for tori implies their 'Minus conjecture', using [50, Chap. IV, Lem. 1.1] to eliminate the dependence on 'T' (just as in the proof of Theorem 6.1). The validity of the 'Minus conjecture' in the case $k=\mathbb{Q}$ is thus also a consequence of Theorem 6.5.

7. Higher Fitting ideals of Selmer groups

In this section, we introduce a natural notion of 'higher *relative* Fitting ideals' in §7.1, and then study the higher Fitting ideals of the transposed Selmer group $\mathcal{S}^{\mathrm{tr}}_{S,T}(\mathbb{G}_{m/K})$. In this way we prove Theorems 1.5 and 1.10 and Corollary 1.14.

7.1. Relative Fitting ideals.

In this subsection, we recall the definition of Fitting ideals and also introduce a natural notion of 'higher *relative* Fitting ideals'.

Suppose that R is a noetherian ring and M is a finitely generated R-module. Take an exact sequence

$$R^{\oplus m}\xrightarrow{f}R^{\oplus n}\to M\to 0,$$

and denote by A_f the matrix with n rows and m columns corresponding to f. Then for $i \in \mathbb{Z}_{\geq 0}$ the i-th Fitting ideal of M, denoted by $\mathrm{Fitt}^i_R(M)$, is defined to be the ideal generated by all $(n-i)\times(n-i)$ minors of A_f if $0 \leq i < n$ and R if $i \geq n$. In this situation we call A_f a *relation matrix* of M. These ideals do not depend on the choice of the above exact sequence (see [40, Chap. 3]). The usual notation is $\mathrm{Fitt}_{i,R}(M)$, but we use the above notation which is consistent with the exterior power $\bigwedge^i_R M$. If we can take a presentation

$$R^{\oplus m} \xrightarrow{f} R^{\oplus n} \to M \to 0$$

of M with $m = n$, then we say M has a *quadratic presentation*.

We now fix a submodule N of M, and non-negative integers a and b. We write ν for the minimal number of generators of N.

If $b > \nu$, then we simply set

$$\mathrm{Fitt}^{(a,b)}_R(M,N) := \mathrm{Fitt}^a_R(M/N).$$

However, if $b \leq \nu$ then we consider a relation matrix for M of the form

$$A = \begin{pmatrix} A_1 & A_2 \\ 0 & A_3 \end{pmatrix}$$

where A_1 is a relation matrix of N. We suppose that A_1 is a matrix with n_1 rows and m_1 columns and A_3 is a matrix with n_2 rows and m_2 columns. We remove b rows from among the first row to the n_1-th row of A to get a matrix A', and remove a rows from A' to get A''. We denote by $F(A'')$ the ideal generated by all $c \times c$ minors of A'' where $c = n_1 + n_2 - a - b$ if $c > 0$ and $F(A'') = R$ otherwise. We consider all such A'' obtained from A and then define the relative Fitting ideal by setting

$$\mathrm{Fitt}^{(a,b)}_R(M,N) := \sum_{A''} F(A'').$$

By the standard method using the elementary operations of matrices (see the proof of [40, p.86, Th. 1]), one can show that this sum does not depend on the choice of relation matrix A.

The following result gives an alternative characterization of this ideal.

Lemma 7.1. *Let X be an R-submodule of M that is generated by $(a+b)$ elements $x_1, \dots, x_{a+b}$ such that the elements $x_1, \dots, x_b$ belong to N. Let $\mathcal{X}$ be the set of such R-submodules of M. Then we have*

$$\mathrm{Fitt}^{(a,b)}_R(M,N) = \sum_{X \in \mathcal{X}} \mathrm{Fitt}^0_R(M/X).$$

Proof. If $b > \nu$, both sides equal $\mathrm{Fitt}^a_R(M/N)$, so we may assume $b \leq \nu$. Let $e_1, \dots, e_n$ be the generators of M corresponding to the above matrix A where $n = n_1 + n_2$. Suppose that A'' is obtained from A by removing $(a+b)$ rows, from the i_1-th row to the i_{a+b}-th row with $1 \leq i_1, \dots, i_b \leq n_1$. Let X be a submodule of M generated by $e_{i_1}, \dots, e_{i_{a+b}}$. Then by definitions $X \in \mathcal{X}$ and

$F(A'') = \mathrm{Fitt}^0_R(M/X)$. This shows that the left hand side of the equation in Lemma 7.1 is in the right hand side.
On the other hand, suppose that X is in $\mathcal{X}$ and $x_1, \ldots, x_{a+b}$ are generators of X. Regarding $e_1, \ldots, e_{n_1}, x_1, \ldots, x_b, e_{n_1+1}, \ldots, e_n, x_{b+1}, \ldots, x_{a+b}$ as generators of M, we have a relation matrix of M of the form

$$B = \begin{pmatrix} A_1 & B_1 & A_2 & B_2 \\ 0 & I_b & 0 & 0 \\ 0 & 0 & A_3 & B_3 \\ 0 & 0 & 0 & I_a \end{pmatrix}$$

where I_a, I_b are the identity matrices of degree a, b, respectively. Then

$$C = \begin{pmatrix} A_1 & B_1 & A_2 & B_2 \\ 0 & 0 & A_3 & B_3 \end{pmatrix}$$

is a relation matrix of M/X. Since C is obtained from B by removing $(a+b)$ rows in the way of obtaining A'' from A, it follows from the definition of the relative Fitting ideal that $\mathrm{Fitt}^0_R(M/X) \subset \mathrm{Fitt}^{(a,b)}_R(M,N)$. This shows the other inclusion. □

In the next result we record some useful properties of higher relative Fitting ideals.

Lemma 7.2.

(i) $\mathrm{Fitt}^{(a,b)}_R(M,N) \subset \mathrm{Fitt}^{a+b}_R(M)$.

(ii) $\mathrm{Fitt}^{(a,0)}_R(M,N) = \mathrm{Fitt}^{a}_R(M)$.

(iii) *Suppose that there exists an exact sequence* $0 \to M' \to M \to R^{\oplus r} \to 0$ *of* R*-modules and that* $N \subset M'$. *Then one has*

$$\mathrm{Fitt}^{(a,b)}_R(M,N) = \begin{cases} \mathrm{Fitt}^{(a-r,b)}_R(M',N), & \text{if } a \geq r, \\ 0, & \text{otherwise.} \end{cases}$$

(iv) *Assume that* M/N *has a quadratic presentation. Then one has*

$$\mathrm{Fitt}^{(0,b)}_R(M,N) = \mathrm{Fitt}^{b}_R(N)\,\mathrm{Fitt}^{0}_R(M/N).$$

Proof. Claims (i), (ii) and (iii) follow directly from the definition of the higher relative Fitting ideal. To prove claim (iv), we consider a relation matrix

$$A = \begin{pmatrix} A_1 & A_2 \\ 0 & A_3 \end{pmatrix}$$

as above, where A_1 is a matrix with n_1 rows and A_3 is a square matrix of n_2 rows. Put $n = n_1 + n_2$. Then a matrix A'' obtained from A as above is of the form

$$A'' = \begin{pmatrix} A''_1 & A''_2 \\ 0 & A_3 \end{pmatrix}.$$

This is a matrix with $(n-b)$ rows and so a nonzero $(n-b) \times (n-b)$ minor of A'' must be $\det(A_3)$ times a $(n_1 - b) \times (n_1 - b)$ minor of A''_1. This implies the required conclusion. □

7.2. Statement of the conjecture. Let $K/k, G, S, T, V$ be as in §5.1. For the element $\epsilon^V_{K/k,S,T}$, the Rubin-Stark conjecture asserts that $\Phi(\epsilon^V_{K/k,S,T})$ belongs to $\mathbb{Z}[G]$ for every Φ in $\bigwedge_G^r \mathrm{Hom}_G(\mathcal{O}^\times_{K,S,T}, \mathbb{Z}[G])$.
We next formulate a much stronger conjecture which describes the arithmetic significance of the ideal generated by the elements $\Phi(\epsilon^V_{K/k,S,T})$ when Φ runs over $\bigwedge_G^r \mathrm{Hom}_G(\mathcal{O}^\times_{K,S,T}, \mathbb{Z}[G])$.

Conjecture 7.3. *One has an equality*

$$\mathrm{Fitt}^r_G(\mathcal{S}_{S,T}(\mathbb{G}_{m/K})) = \{\Phi(\epsilon^V_{K/k,S,T})^{\#} : \Phi \in \bigwedge_G^r \mathrm{Hom}_G(\mathcal{O}^\times_{K,S,T}, \mathbb{Z}[G])\},$$

or equivalently (by Lemma 2.8),

$$\mathrm{Fitt}^r_G(\mathcal{S}^{\mathrm{tr}}_{S,T}(\mathbb{G}_{m/K})) = \{\Phi(\epsilon^V_{K/k,S,T}) : \Phi \in \bigwedge_G^r \mathrm{Hom}_G(\mathcal{O}^\times_{K,S,T}, \mathbb{Z}[G])\}.$$

The following result shows that this conjecture refines the first half of the statement of Conjecture 5.9.

Proposition 7.4. *For a finite set Σ of places, we put $\mathcal{J}_\Sigma = \prod_{v\in\Sigma} I(G_v)\mathbb{Z}[G]$. Assume Conjecture 7.3 is valid. Then, for every $\Phi \in \bigwedge_G^r \mathrm{Hom}_G(\mathcal{O}^\times_{K,S,T}, \mathbb{Z}[G])$ and $v \in S \setminus V$, one has*

$$\Phi(\epsilon^V_{K/k,S,T}) \in \mathcal{J}_{S\setminus(V\cup\{v\})}.$$

Proof. It is sufficient to show that $\mathrm{Fitt}^r_G(\mathcal{S}^{\mathrm{tr}}_{S,T}(\mathbb{G}_{m/K})) \subset \mathcal{J}_{S\setminus(V\cup\{v\})}$. Since there is a canonical surjective homomorphism

$$\mathcal{S}^{\mathrm{tr}}_{S,T}(\mathbb{G}_{m/K}) \longrightarrow X_{K,S} \simeq \mathbb{Z}[G]^{\oplus r} \oplus X_{K,S\setminus V},$$

we have

$$\mathrm{Fitt}^r_G(\mathcal{S}^{\mathrm{tr}}_{S,T}(\mathbb{G}_{m/K})) \subset \mathrm{Fitt}^r_G(X_{K,S}) = \mathrm{Fitt}^0_G(X_{K,S\setminus V}).$$

The existence of the surjective homomorphism $X_{K,S\setminus V} \to Y_{K,S\setminus(V\cup\{v\})}$ implies that $\mathrm{Fitt}^0_G(X_{K,S\setminus V}) \subset \mathrm{Fitt}^0_G(Y_{K,S\setminus(V\cup\{v\})}) = \mathcal{J}_{S\setminus(V\cup\{v\})}$. This completes the proof. □

7.3. The leading term conjecture implies Conjecture 7.3. The following result combines with Lemma 2.8 to imply the statement of Theorem 1.5(i).

Theorem 7.5. LTC(K/k) *implies Conjecture 7.3. In particular, Conjecture 7.3 is valid if either K is an abelian extension over $\mathbb{Q}$ or k is a function field or $[K : k] \leq 2$.*

Proof. The second claim is a consequence of Remark 3.3.
To prove the first claim we assume the validity of LTC(K/k). Then the module P that occurs in the exact sequence (24) is free of rank d, as we noted before.

Hence we may assume $P=F$. Let $z_b \in \bigwedge_G^d F$ be as in §5.6. By LTC(K/k), z_b is a G-basis of $\bigwedge_G^d F$. Write z_b as

$$z_b = x \bigwedge_{1\le i\le d} b_i$$

with some $x \in \mathbb{Z}[G]^\times$. By Theorem 5.14 and Proposition 4.1, we have

$$\epsilon^V_{K/k,S,T} = \pm x \sum_{\sigma\in\mathfrak{S}_{d,r}} \operatorname{sgn}(\sigma)\det(\psi_i(b_{\sigma(j)}))_{r<i,j\le d} b_{\sigma(1)}\wedge\cdots\wedge b_{\sigma(r)}.$$

Take $\Phi \in \bigwedge_G^r \operatorname{Hom}_G(\mathcal{O}^\times_{K,S,T},\mathbb{Z}[G])$. Since $F/\mathcal{O}^\times_{K,S,T}\simeq \operatorname{im}(\psi)\subset F$ is $\mathbb{Z}$-torsion-free, we know by Lemma 4.7(ii) that the map

$$\operatorname{Hom}_G(F,\mathbb{Z}[G]) \longrightarrow \operatorname{Hom}_G(\mathcal{O}^\times_{K,S,T},\mathbb{Z}[G])$$

induced by the inclusion $\mathcal{O}^\times_{K,S,T}\subset F$ is surjective. Hence, we can take a lift $\widetilde{\Phi}$ of Φ to $\bigwedge_G^r \operatorname{Hom}_G(F,\mathbb{Z}[G])$. We have

$$\begin{aligned}\Phi(\epsilon^V_{K/k,S,T}) &= \pm x \sum_{\sigma\in\mathfrak{S}_{d,r}} \operatorname{sgn}(\sigma)\det(\psi_i(b_{\sigma(j)}))_{r<i,j\le d}\widetilde{\Phi}(b_{\sigma(1)}\wedge\cdots\wedge b_{\sigma(r)})\\ &\in \langle \det(\psi_i(b_{\sigma(j)}))_{r<i,j\le d} : \sigma\in\mathfrak{S}_{d,r}\rangle_G.\end{aligned}$$

We consider the matrix A corresponding to the presentation

$$F\to F\to \mathcal{S}^{\mathrm{tr}}_{S,T}(\mathbb{G}_{m/K})\to 0$$

which comes from the exact sequence (24). By Lemma 5.20, $\psi_i=0$ for $1\le i\le r$. If we write elements in F as column vectors, this implies that the i-th row of A is zero for all i such that $1\le i\le r$. Hence we have

$$\operatorname{Fitt}^r_G(\mathcal{S}^{\mathrm{tr}}_{S,T}(\mathbb{G}_{m/K})) = \langle \det(\psi_i(b_{\sigma(j)}))_{r<i,j\le d} : \sigma\in\mathfrak{S}_{d,r}\rangle_G.$$

Therefore, we get an inclusion

$$\{\Phi(\epsilon^V_{K/k,S,T}) : \Phi\in \bigwedge_G^r \operatorname{Hom}_G(\mathcal{O}^\times_{K,S,T},\mathbb{Z}[G])\}\subset \operatorname{Fitt}^r_G(\mathcal{S}^{\mathrm{tr}}_{S,T}(\mathbb{G}_{m/K})).$$

We obtain the reverse inclusion from

$$(b^*_{\sigma(1)}\wedge\cdots\wedge b^*_{\sigma(r)})(\epsilon^V_{K/k,S,T}) = \pm x\det(\psi_i(b_{\sigma(j)}))_{r<i,j\le d}$$

and the fact that x is a unit in $\mathbb{Z}[G]$. □

7.4. The proof of Theorem 1.10. For any G-module M we write M^* for the linear dual $\operatorname{Hom}_{\mathbb{Z}}(M,\mathbb{Z})$ endowed with the natural contragredient action of G. We also set $V' := V\cup\{v\}$.

We start with a useful technical observation.

Lemma 7.6. *For each integer i with $1\le i\le r$ let φ_i be an element of $(\mathcal{O}^\times_{K,S,T})^*$. Then for any given integer N there is a subset $\{\varphi'_i : 1\le i\le r\}$ of $(\mathcal{O}^\times_{K,S,T})^*$ which satisfies the following properties.*

(i) *For each i one has $\varphi'_i\equiv\varphi_i$ modulo $N\cdot(\mathcal{O}^\times_{K,S,T})^*$.*

(ii) *The image in* $(\mathcal{O}_{K,V',T}^{\times})^*$ *of the submodule of* $(\mathcal{O}_{K,S,T}^{\times})^*$ *that is generated by the set* $\{\varphi_i' : 1 \le i \le r\}$ *is free of rank* r.

Proof. Our choice of V implies that we may choose a free G-submodule $\mathcal{F}$ of $(\mathcal{O}_{K,V',T}^{\times})^*$ of rank r. We then choose a subset $\{f_i : 1 \le i \le r\}$ of $(\mathcal{O}_{K,S,T}^{\times})^*$ which the natural surjection $\rho : (\mathcal{O}_{K,S,T}^{\times})^* \to (\mathcal{O}_{K,V',T}^{\times})^*$ sends to a basis of $\mathcal{F}$. For any integer m we set $\varphi_{i,m} := \varphi_i + mNf_i$ and note it suffices to show that for any sufficiently large m the elements $\{\rho(\varphi_{i,m}) : 1 \le i \le r\}$ are linearly independent over $\mathbb{Q}[G]$.

Consider the composite homomorphism of G-modules $\mathcal{F} \to \mathbb{Q}(\mathcal{O}_{K,V',T}^{\times})^* \to \mathbb{Q}\mathcal{F}$ where the first arrow sends each $\rho(f_i)$ to $\rho(\varphi_{i,m})$ and the second is induced by a choice of $\mathbb{Q}[G]$-equivariant section to the projection $\mathbb{Q}(\mathcal{O}_{K,V',T}^{\times})^* \to \mathbb{Q}((\mathcal{O}_{K,V',T}^{\times})^*/\mathcal{F})$. Then, with respect to the basis $\{\rho(f_i) : 1 \le i \le r\}$, this linear map is represented by a matrix of the form $A + mNI_r$ for a matrix A in $\mathrm{M}_r(\mathbb{Q}[G])$ that is independent of m. In particular, if m is large enough to ensure that $-mN$ is not an eigenvalue of $e_\chi A$ for any χ, then the composite homomorphism is injective and so the elements $\{\rho(\varphi_{i,m}) : 1 \le i \le r\}$ are linearly independent over $\mathbb{Q}[G]$, as required. □

For each integer i with $1 \le i \le r$ let φ_i be an element of $(\mathcal{O}_{K,S,T}^{\times})^*$. Then, for any non-zero integer N which belongs to $\mathrm{Fitt}_G^0(\mathrm{Cl}(K))$ we choose homomorphisms φ_i' as in Lemma 7.6. Then the congruences in Lemma 7.6(i) imply that

$$(\bigwedge_{1\le i\le r} \varphi_i)(\epsilon_{K/k,S,T}^V) \equiv (\bigwedge_{1\le i\le r} \varphi_i')(\epsilon_{K/k,S,T}^V) \text{ modulo } N\cdot\mathbb{Z}[G].$$

Given this, Lemma 7.6(ii) implies that Theorem 1.10 is true provided that it is true for all Φ of the form $\bigwedge_{1\le i\le r}\varphi_i$ where the images in $(\mathcal{O}_{K,V',T}^{\times})^*$ of the homomorphisms φ_i span a free module of rank r.

We shall therefore assume in the sequel that Φ is of this form.

For each index i we now choose a lift $\widetilde{\varphi}_i$ of φ_i to $\mathcal{S}_{S,T}(\mathbb{G}_{m/K})$ and then write $\mathcal{E}_\Phi$ for the G-module that is generated by the set $\{\widetilde{\varphi}_i : 1 \le i \le r\}$.

Proposition 7.7. *If* $\mathrm{LTC}(K/k)$ *is valid, then for every* Φ *as above one has*

$$\Phi(\epsilon_{K/k,S,T}^V)^{\#} \in \mathrm{Fitt}_G^0(\mathcal{S}_{S,T}(\mathbb{G}_{m/K})/\mathcal{E}_\Phi).$$

Proof. We use the existence of an exact triangle in $D^{\mathrm{p}}(\mathbb{Z}[G])$ of the form

$$\mathbb{Z}[G]^{\oplus r,\bullet} \xrightarrow{\theta} R\Gamma_{c,T}((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{Z}) \xrightarrow{\theta'} C^\bullet \to \mathbb{Z}[G]^{\oplus r,\bullet}[1]. \tag{29}$$

Here $\mathbb{Z}[G]^{\oplus r,\bullet}$ denotes the complex $\mathbb{Z}[G]^{\oplus r}[-1] \oplus \mathbb{Z}[G]^{\oplus r}[-2]$ and, after choosing an ordering $\{v_i : 1 \le i \le r\}$ of the places in V, the morphism θ is uniquely specified by the condition that $H^1(\theta)$ sends each element b_i of the canonical basis $\{b_i : 1 \le i \le r\}$ of $\mathbb{Z}[G]^{\oplus r}$ to w_i^* in $(Y_{K,V})^* \subset (X_{K,S})^* = H^1_{c,T}((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{Z})$ and $H^2(\theta)$ sends each b_i to $\widetilde{\varphi}_i$ in $\mathcal{S}_{S,T}(\mathbb{G}_{m/K})$.

Note that the long exact cohomology sequence of this triangle implies $C^\bullet$ is acyclic outside degrees one and two and identifies $H^1(C^\bullet)$ and $H^2(C^\bullet)$ with $(X_{K,S\setminus V})^*$ and $\mathcal{S}_{S,T}(\mathbb{G}_{m/K})/\mathcal{E}_\Phi$, respectively.

In particular, if we now write e_r for the idempotent of $\mathbb{Q}[G]$ obtained as $\sum_{r_{\chi,S}=r} e_\chi$, then the space $e_r\mathbb{Q}H^i(C^\bullet)$ vanishes for both $i=1$ and $i=2$. We may therefore choose a commutative diagram of $\mathbb{R}[G]$-modules

(30)

$$\begin{CD}
0 @>>> \mathbb{R}[G]^{\oplus r} @>{H^1(\theta)}>> \mathbb{R}H^1_{c,T}((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{Z}) @>{H^1(\theta')}>> \mathbb{R}H^1(C^\bullet) @>>> 0\\
@. @V{\lambda_1}VV @V{\lambda_2}VV @V{\lambda_3}VV @.\\
0 @>>> \mathbb{R}[G]^{\oplus r} @>{H^2(\theta)}>> \mathbb{R}H^2_{c,T}((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{Z}) @>{H^2(\theta')}>> \mathbb{R}H^2(C^\bullet) @>>> 0
\end{CD}$$

such that $e_r\lambda_2 = e_r\lambda^*_{K,S}$.

This diagram combines with the triangle (29) to imply that there is an equality of lattices

$$\vartheta_{\lambda_2}(\mathrm{det}_G(R\Gamma_{c,T}((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{Z})))^{-1} = \det(\lambda_1)\cdot\vartheta_{\lambda_3}(\mathrm{det}_G(C^\bullet))^{-1}. \tag{31}$$

We now assume that the conjecture LTC(K/k) is valid. Then Proposition 3.4 implies that $\mathrm{det}_G(R\Gamma_{c,T}((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{Z}))^{-1}$ is a free rank one $\mathbb{Z}[G]$-module and further that if we choose any basis ξ for this module, then both $e_r\xi$ and $e_r\theta^*_{K/k,S,T}(0)^\# = \theta^{(r),\#}_{K/k,S,T}$ are bases of the $e_r\mathbb{Z}[G]$-module

$$e_r\vartheta_{\lambda_2}(\mathrm{det}_G(R\Gamma_{c,T}((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{Z})))^{-1} = e_r\vartheta_{\lambda^*_{K,S}}(\mathrm{det}_G(R\Gamma_{c,T}((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{Z})))^{-1}.$$

Bass's Theorem (cf. [32, Chap. 7, (20.9)]) implies that for each prime p the projection map $\mathbb{Z}_{(p)}[G]^\times \to e_r\mathbb{Z}_{(p)}[G]^\times$ is surjective. The above equality thus implies that the $\mathbb{Z}_{(p)}[G]$-module $\vartheta_{\lambda_2}(\mathrm{det}_G(R\Gamma_{c,T}((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{Z})))^{-1}\otimes_{\mathbb{Z}}\mathbb{Z}_{(p)}$ has a basis ξ_p for which one has $e_r\xi_p = e_r\theta^*_{K/k,S,T}(0)^\# = \theta^{(r),\#}_{K/k,S,T}$. For each prime p the equality (31) therefore implies that

$$\begin{aligned}
&e_r\vartheta_{\lambda_3}(\mathrm{det}_G(C^\bullet))^{-1}\otimes_{\mathbb{Z}}\mathbb{Z}_{(p)}\\
=&\det(\lambda_1)^{-1}e_r\vartheta_{\lambda_2}(\mathrm{det}_G(R\Gamma_{c,T}((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{Z})))^{-1}\otimes_{\mathbb{Z}}\mathbb{Z}_{(p)}\\
=&\mathbb{Z}_{(p)}[G]\cdot e_r\det(\lambda_1)^{-1}\theta^{(r),\#}_{K/k,S,T}.
\end{aligned} \tag{32}$$

Now the commutativity of (30) implies that $e_r\det(\lambda_1)$ is equal to the determinant of the matrix which represents $e_r\lambda^*_{K,S}$ with respect to the bases $\{e_rw^*_i : 1\le i\le r\}$ and $\{e_r\varphi_i : 1\le i\le r\}$ and hence that

$$e_r\bigwedge_{1\le i\le r}\lambda^*_{K,S}(w^*_i) = e_r\det(\lambda_1)\Phi.$$

Since the element $\epsilon^V_{K/k,S,T}$ is defined via the equality

$$\theta^{(r)}_{K/k,S,T}\bigwedge_{1\le i\le r}\lambda^{-1}_{K,S}(w_i - w) = \epsilon^V_{K/k,S,T},$$

one therefore has

$$\begin{aligned}
&\Phi(\epsilon^V_{K/k,S,T})^\# \\
(33) \qquad &= (e_r\mathrm{det}(\lambda_1))^{-1} \bigwedge_{1\le i\le r} \lambda^*_{K,S}(w_i^*)(\theta^{(r),\#}_{K/k,S,T}(\bigwedge_{1\le i\le r} \lambda^{-1}_{K,S}(w_i - w))) \\
&= (e_r\mathrm{det}(\lambda_1))^{-1}\theta^{(r),\#}_{K/k,S,T} \\
&\in e_r\vartheta_{\lambda_3}(\mathrm{det}_G(C^\bullet))^{-1}\otimes_{\mathbb{Z}}\mathbb{Z}_{(p)}
\end{aligned}$$

where the last containment follows from (32).
Now by the same reasoning as used in the proof of Lemma 2.8, we know that the p-localized complex $\mathbb{Z}_{(p)}\otimes C^\bullet$ is represented by a complex $P \xrightarrow{\delta} P$, where P is a finitely generated free $\mathbb{Z}_{(p)}[G]$-module and the first term is placed in degree one. In particular, since for any character χ of G the space $e_\chi\mathbb{C}H^1(C^\bullet) = e_\chi\mathbb{C}\ker(\delta)$ does not vanish if $e_\chi e_r = 0$, one has

$$\begin{aligned}
(34) \qquad e_r\vartheta_{\lambda_3}(\mathrm{det}_G(C^\bullet))^{-1}_{(p)} &= \mathbb{Z}_{(p)}[G]e_r\mathrm{det}(\delta) \\
&= \mathbb{Z}_{(p)}[G]\mathrm{det}(\delta) \\
&\subset \mathrm{Fitt}^0_{\mathbb{Z}_{(p)}[G]}((H^2_{c,T}((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{Z})/\mathcal{E}_\Phi)\otimes_{\mathbb{Z}}\mathbb{Z}_{(p)}) \\
&= \mathrm{Fitt}^0_G(H^2_{c,T}((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{Z})/\mathcal{E}_\Phi)\otimes_{\mathbb{Z}}\mathbb{Z}_{(p)}.
\end{aligned}$$

The inclusion here follows from the tautological exact sequence

$$P \xrightarrow{\delta} P \longrightarrow H^2(\mathbb{Z}_{(p)}\otimes C^\bullet) \longrightarrow 0$$

and the identification $H^2(\mathbb{Z}_{(p)}\otimes C^\bullet) = (H^2_{c,T}((\mathcal{O}_{K,S})_{\mathcal{W}},\mathbb{Z})/\mathcal{E}_\Phi)\otimes_{\mathbb{Z}}\mathbb{Z}_{(p)}$.
The claimed result now follows from (33) and (34). □

Now we proceed to the proof of Theorem 1.10. The existence of a surjective homomorphism of G-modules $f : \mathcal{S}_{S,T}(\mathbb{G}_{m/K}) \to \mathcal{S}_{V'\cup S_\infty,T}(\mathbb{G}_{m/K})$ (see Proposition 2.4(ii)) combines with Proposition 7.7 to imply that

$$(35) \qquad \Phi(\epsilon^V_{K/k,S,T})^\# \in \mathrm{Fitt}^0_G(\mathcal{S}_{V'\cup S_\infty,T}(\mathbb{G}_{m/K})/f(\mathcal{E}_\Phi)).$$

This implies the first assertion of Theorem 1.10 since the natural map $\mathrm{Cl}^T_{V'}(K)^\vee \to \mathcal{S}_{V'\cup S_\infty,T}(\mathbb{G}_{m/K})$ induces an injection

$$\mathrm{Cl}^T_{V'}(K)^\vee \to \mathcal{S}_{V'\cup S_\infty,T}(\mathbb{G}_{m/K})/f(\mathcal{E}_\Phi).$$

In addition, if G is cyclic, then the latter injection combines with (35) to imply that

$$\Phi(\epsilon^V_{K/k,S,T}) \in \mathrm{Fitt}^0_G(\mathrm{Cl}^T_{V'}(K)^\vee)^\# = \mathrm{Fitt}^0_G(\mathrm{Cl}^T_{V'}(K)),$$

as claimed by the second assertion of Theorem 1.10.
This completes the proof of Theorem 1.10.

7.5. The proof of Corollary 1.14. Let K/k be a CM-extension, $S = S_\infty(k)$, and p an odd prime. For a $\mathbb{Z}_p[G]$-module M, we denote by M^- the submodule on which the complex conjugation acts as -1.
Then, since complex conjugation acts trivially on $\mathrm{Hom}_{\mathbb{Z}}(\mathcal{O}_{K,S,T}^\times, \mathbb{Z}) \otimes \mathbb{Z}_p$, the exact sequence

$$0 \longrightarrow \mathrm{Cl}^T(K)^\vee \longrightarrow \mathcal{S}_{S,T}(\mathbb{G}_{m/K}) \longrightarrow \mathrm{Hom}_{\mathbb{Z}}(\mathcal{O}_{K,S,T}^\times, \mathbb{Z}) \longrightarrow 0,$$

implies that in this case there is an equality

$$((\mathrm{Cl}^T(K) \otimes \mathbb{Z}_p)^\vee)^- = (\mathcal{S}_{S,T}(\mathbb{G}_{m/K}) \otimes \mathbb{Z}_p)^-.$$

In addition, in this case the containment of Proposition 7.7 applies with V empty (so $r = 0$ and $\mathcal{E}_\Phi$ vanishes) to imply that

$$\theta_{K/k,S,T}(0)^\# \in \mathrm{Fitt}_G^0(\mathcal{S}_{S,T}(\mathbb{G}_{m/K})),$$

and hence one has

$$\theta_{K/k,S,T}(0)^\# \in \mathrm{Fitt}^0_{\mathbb{Z}_p[G]}(((\mathrm{Cl}^T(K) \otimes \mathbb{Z}_p)^\vee)^-).$$

Since $\theta_{K/k,S,T}(0)$ lies in the minus component of $\mathbb{Z}_p[G]$, this is in turn equivalent to the required containment

$$\theta_{K/k,S,T}(0)^\# \in \mathrm{Fitt}^0_{\mathbb{Z}_p[G]}((\mathrm{Cl}^T(K) \otimes \mathbb{Z}_p)^\vee).$$

This completes the proof of Corollary 1.14.

7.6. The higher relative Fitting ideals of the dual Selmer group. We write M_{tors} for the $\mathbb{Z}$-torsion submodule of a G-module M and abbreviate the higher relative Fitting ideal $\mathrm{Fitt}^{(a,b)}_{\mathbb{Z}[G]}(M, M_{\mathrm{tors}})$ to $\mathrm{Fitt}^{(a,b)}_G(M)$.
In this subsection, we study the ideals $\mathrm{Fitt}^{(r,i)}_G(\mathcal{S}^{\mathrm{tr}}_{S,T}(\mathbb{G}_{m/K}))$ and, in particular, prove Theorem 1.5(ii). We note that the exact sequence (2) identifies $\mathcal{S}^{\mathrm{tr}}_{S,T}(\mathbb{G}_{m/K})_{\mathrm{tors}}$ with the group $\mathrm{Cl}^T_S(K)$.
For each non-negative integer i we define the set $\mathcal{V}_i$ as in Theorem 1.5(ii).

Conjecture 7.8. *For each non-negative integer i one has an equality*

$$\begin{aligned}&\mathrm{Fitt}^{(r,i)}_G(\mathcal{S}^{\mathrm{tr}}_{S,T}(\mathbb{G}_{m/K}))\\ &= \{\Phi(\epsilon^{V\cup V'}_{K/k,S\cup V',T}) : V' \in \mathcal{V}_i \text{ and } \Phi \in \bigwedge_G^{r+i} \mathrm{Hom}_G(\mathcal{O}^\times_{K,S\cup V',T}, \mathbb{Z}[G])\}.\end{aligned}$$

The following result is a generalization of Theorem 7.5 in §7.3.

Theorem 7.9. *If* $\mathrm{LTC}(K/k)$ *is valid, then so is Conjecture 7.8.*

Proof. We consider the composition of the two canonical homomorphisms

$$\mathcal{S}^{\mathrm{tr}}_{S,T}(\mathbb{G}_{m/K}) \to X_{K,S} \to Y_{K,V},$$

and denote its kernel by $\mathcal{S}^{\mathrm{tr}}_{S,T}(\mathbb{G}_{m/K})'$. By Lemma 7.2 (iii), we have

$$\mathrm{Fitt}^{(r,i)}_G(\mathcal{S}^{\mathrm{tr}}_{S,T}(\mathbb{G}_{m/K})) = \mathrm{Fitt}^{(0,i)}_G(\mathcal{S}^{\mathrm{tr}}_{S,T}(\mathbb{G}_{m/K})'). \tag{36}$$

We also note that the sequence (2) gives rise to an exact sequence of G-modules

$$0 \longrightarrow \mathrm{Cl}_S^T(K) \longrightarrow \mathcal{S}^{\mathrm{tr}}_{S,T}(\mathbb{G}_{m/K})' \longrightarrow X_{K,S\setminus V} \longrightarrow 0. \tag{37}$$

For $V' \in \mathcal{V}_i$, we denote by $\mathcal{S}^{\mathrm{tr}}_{S\cup V',T}(\mathbb{G}_{m/K})'$ the kernel of the natural composition

$$\mathcal{S}^{\mathrm{tr}}_{S\cup V',T}(\mathbb{G}_{m/K}) \to X_{K,S\cup V'} \to Y_{K,V\cup V'}$$

so that the following sequence is exact

$$0 \longrightarrow \mathrm{Cl}_{S\cup V'}^T(K) \longrightarrow \mathcal{S}^{\mathrm{tr}}_{S\cup V',T}(\mathbb{G}_{m/K})' \longrightarrow X_{K,S\setminus V} \longrightarrow 0.$$

Let $X_{V'}$ be the subgroup of $\mathrm{Cl}_S^T(K)$ generated by the classes of places of K above V' in $\mathrm{Cl}_S^T(K)$. Since $\mathrm{Cl}_S^T(K)/X_{V'} = \mathrm{Cl}_{S\cup V'}^T(K)$, there is an isomorphism $\mathcal{S}^{\mathrm{tr}}_{S,T}(\mathbb{G}_{m/K})'/X_{V'} \simeq \mathcal{S}^{\mathrm{tr}}_{S\cup V',T}(\mathbb{G}_{m/K})'$. By Chebotarev density theorem and Lemma 7.1, we obtain

$$\begin{aligned} \mathrm{Fitt}_G^{(0,i)}(\mathcal{S}^{\mathrm{tr}}_{S,T}(\mathbb{G}_{m/K})') &= \sum_{V'\in\mathcal{V}'_i} \mathrm{Fitt}_G^0(\mathcal{S}^{\mathrm{tr}}_{S\cup V',T}(\mathbb{G}_{m/K})') \\ &= \sum_{V'\in\mathcal{V}'_i} \mathrm{Fitt}_G^{r+i}(\mathcal{S}^{\mathrm{tr}}_{S\cup V',T}(\mathbb{G}_{m/K})) \end{aligned} \tag{38}$$

where we used Lemma 7.2 (iii) again to get the last equality.
Now Theorem 7.9 follows from (36), (38) and Theorem 7.5. □

Corollary 7.10. *We assume that* LTC(K/k) *is valid and that the group* $G = \mathrm{Gal}(K/k)$ *is cyclic. Then for each non-negative integer i one has an equality*

$$\begin{aligned} &\mathrm{Fitt}_G^i(\mathrm{Cl}_S^T(K))\,\mathrm{Fitt}_G^0(X_{K,S\setminus V}) \\ &\quad = \{\Phi(\epsilon^{V\cup V'}_{K/k,S\cup V',T}) : V' \in \mathcal{V}_i \text{ and } \Phi \in {\textstyle\bigwedge_G^{r+i}} \mathrm{Hom}_G(\mathcal{O}^\times_{K,S\cup V',T}, \mathbb{Z}[G])\}. \end{aligned}$$

Proof. Since G is cyclic, the G-module $X_{K,S\setminus V}$ has a quadratic presentation. We may therefore apply Lemma 7.2(iv) to the exact sequence (37) to obtain an equality

$$\mathrm{Fitt}_G^i(\mathrm{Cl}_S^T(K))\,\mathrm{Fitt}_G^0(X_{K,S\setminus V}) = \mathrm{Fitt}_G^{(0,i)}(\mathcal{S}^{\mathrm{tr}}_{S,T}(\mathbb{G}_{m/K})').$$

Given this equality, the claimed result follows from Theorem 7.9 and the equality (36). □

An application of Theorem 7.9 to character components of ideal class groups will be given in §8.

8. Higher Fitting ideals of character components of class groups

In this section, as an application of Theorem 7.9, we study the higher Fitting ideals of character components of class groups.

8.1. General abelian extensions. We suppose that K/k is a finite abelian extension as in §7. We take and fix an odd prime p in this section. We put $A_S^T(K) = \mathrm{Cl}_S^T(K) \otimes \mathbb{Z}_p$, $A^T(K) = \mathrm{Cl}^T(K) \otimes \mathbb{Z}_p$, and $A(K) = \mathrm{Cl}(K) \otimes \mathbb{Z}_p$.
We take a character χ of $G = \mathrm{Gal}(K/k)$. Throughout this section, we assume that *the order of* χ *is prime to* p.
We decompose $G = \Delta_K \times \Gamma_K$ where $|\Delta_K|$ is prime to p and Γ_K is a p-group. By our assumption χ is regarded as a character of Δ_K. For any $\mathbb{Z}_p[\Delta_K]$-module M, we define the χ-component M^χ by setting

$$M^\chi := M \otimes_{\mathbb{Z}_p[\Delta_K]} \mathcal{O}_\chi$$

where $\mathcal{O}_\chi = \mathbb{Z}_p[\mathrm{im}(\chi)]$ on which Δ_K acts via χ. This is an exact functor from the category of $\mathbb{Z}_p[G]$-modules to that of $\mathcal{O}_\chi[\Gamma_K]$-modules.
Let k_χ be the subfield of K corresponding to the kernel of χ, namely, χ induces a faithful character of $\mathrm{Gal}(k_\chi/k)$. Put $K_{(\Delta)} := K^{\Gamma_K}$, then $k_\chi \subset K_{(\Delta)}$. We also put $\Delta_{K,\chi} := \mathrm{Gal}(K_{(\Delta)}/k_\chi)$ which is a subgroup of Δ_K. We consider $K(\chi) := K^{\Delta_{K,\chi}}$, then $\mathrm{Gal}(K(\chi)/k_\chi) = \Gamma_K$. We consider $A_S^T(K)^\chi$ which we regard as an $\mathcal{O}_\chi[\Gamma_K]$-module. By the standard norm argument, we know the natural map $A_S^T(K(\chi))^\chi \to A_S^T(K)^\chi$ is bijective, so when we consider the χ-component $A_S^T(K)^\chi$, we may assume that χ is a faithful character of Δ_K by replacing K with $K(\chi)$. In the following, we assume this. We write $\chi(v) \neq 1$ if the decomposition group of Δ_K at v is non-trivial.
We denote the χ-component of $\epsilon^V_{K/k,S,T}$ by $\epsilon^{V,\chi}_{K/k,S,T} \in ((\bigcap_G^r \mathcal{O}^\times_{K,S,T}) \otimes \mathbb{Z}_p)^\chi$. Let $\mathcal{V}_i$ be the set as in Theorem 1.5(ii) for $i \geq 0$.
Finally we assume that the following condition is satisfied

(∗) any ramifying place v of k in K does not split completely in $K_{(\Delta)}$.

Theorem 8.1. *Let V be the set of the archimedean places of k that split completely in K and set $r := |V|$. We assume that $\chi \neq 1$ is a faithful character of Δ_K, and consider the χ-component of the class group $A^T(K)^\chi$ which is an $\mathcal{O}_\chi[\Gamma_K]$-module. We assume that the χ-component of* $\mathrm{LTC}(K/k)$ *is valid and that the condition* (∗) *is satisfied.*
Then for any non-negative integer i one has an equality

$$\mathrm{Fitt}^i_{\mathcal{O}_\chi[\Gamma_K]}(A^T(K)^\chi) = \{\Phi(\epsilon^{V\cup V',\chi}_{K/k,S\cup V',T}) : V' \in \mathcal{V}_i \text{ and } \Phi \in \bigwedge^{r+i}_{\mathcal{O}_\chi[\Gamma_K]} \mathcal{H}_\chi\}$$

where $S = S_\infty(k) \cup S_{\mathrm{ram}}(K/k)$ *and* $\mathcal{H}_\chi = \mathrm{Hom}_{\mathcal{O}_\chi[\Gamma_K]}((\mathcal{O}^\times_{K,S\cup V',T} \otimes \mathbb{Z}_p)^\chi, \mathcal{O}_\chi[\Gamma_K])$.

Proof. Since $v \in S_{\mathrm{ram}}(K/k)$ does not split completely in $K_{(\Delta)}$, one has $\chi(v) \neq 1$ and hence $(Y_{K,S_{\mathrm{ram}}} \otimes \mathbb{Z}_p)^\chi = 0$.
As $\chi \neq 1$, we therefore also have $(X_{K,S_{\mathrm{ram}}} \otimes \mathbb{Z}_p)^\chi = (Y_{K,S_{\mathrm{ram}}} \otimes \mathbb{Z}_p)^\chi = 0$. Hence $(X_{K,S} \otimes \mathbb{Z}_p)^\chi = (Y_{K,S_\infty} \otimes \mathbb{Z}_p)^\chi$ is isomorphic to $\mathcal{O}_\chi[\Gamma_K]^{\oplus r}$. This implies that

$$\mathrm{Fitt}^{(r,i)}_{\mathcal{O}_\chi[\Gamma_K]}((\mathcal{S}^{\mathrm{tr}}_{S,T}(\mathbb{G}_{m/K}) \otimes \mathbb{Z}_p)^\chi), A_S^T(K)^\chi) = \mathrm{Fitt}^i_{\mathcal{O}_\chi[\Gamma_K]}(A_S^T(K)^\chi)$$

and so the claim follows from Theorem 7.9. □

In the case $K = k_\chi$, the condition $(*)$ is automatically satisfied. We denote the group $A^T(k_\chi)^\chi$ by $(A^T)^\chi$, which is determined only by χ.

Corollary 8.2. *Let χ be a non-trivial linear character of k of order prime to p, and V the set of the archimedean places of k that split completely in k_χ and set $r := |V|$. We assume the χ-component of* $\mathrm{LTC}(k_\chi/k)$ *to be valid. Then for any non-negative integer i one has an equality*

$$\mathrm{Fitt}^i_{\mathcal{O}_\chi}((A^T)^\chi) = \{\Phi(\epsilon^{V\cup V',\chi}_{k_\chi/k,S\cup V',T}) : V' \in \mathcal{V}_i \text{ and } \Phi \in \bigwedge^{r+i}_{\mathcal{O}_\chi} \mathcal{H}_\chi\}$$

where $S = S_\infty(k) \cup S_{\mathrm{ram}}(k_\chi/k)$ and $\mathcal{H}_\chi = \mathrm{Hom}_{\mathcal{O}_\chi}((\mathcal{O}^\times_{k_\chi,S\cup V',T} \otimes \mathbb{Z}_p)^\chi, \mathcal{O}_\chi)$.

8.2. The order of character components in CM abelian extensions. In this subsection, we assume that k is totally real, K is a CM-field, and χ is an odd character. In this case, we can compute the right hand side of Theorem 8.1 more explicitly. First of all, note that $r = 0$ in this case.

We first consider the case $K = k_\chi$ and $i = 0$. When $S = S_\infty(k) \cup S_{\mathrm{ram}}(k_\chi/k)$, we denote the L-function $L_{k,S,T}(\chi^{-1}, s)$ by $L^T_k(\chi^{-1}, s)$. When T is empty, we denote $L^T_k(\chi^{-1}, s)$ by $L_k(\chi^{-1}, s)$. In this case, we know

$$\epsilon^{\emptyset,\chi}_{k_\chi/k,S,T} = \theta_{k_\chi/k,S,T}(0)^\chi = L^T_k(\chi^{-1}, 0)$$

(see §5.3). Therefore, Corollary 8.2 with $i = 0$ implies

Corollary 8.3. *Let k be totally real, and χ a one dimensional odd character of k of order prime to p. We assume the χ-component of* $\mathrm{LTC}(k_\chi/k)$ *to be valid.*

(i) *One has $|(A^T)^\chi| = |\mathcal{O}_\chi/L^T_k(\chi^{-1}, 0)|$.*

(ii) *Let $\mathrm{Cl}(k_\chi)$ be the ideal class group of k_χ, $A(k_\chi) = \mathrm{Cl}(k_\chi) \otimes \mathbb{Z}_p$, and $A^\chi = A(k_\chi)^\chi$. We denote by ω the Teichmüller character giving the Galois action on μ_p, the group of p-th roots of unity, and by $\mu_{p^\infty}(k(\mu_p))$ the group of roots of unity of p-power order in $k(\mu_p)$. Then one has*

$$|A^\chi| = \begin{cases} |\mathcal{O}_\chi/L_k(\chi^{-1}, 0)| & \text{if } \chi \neq \omega, \\ |\mathcal{O}_\chi/(|\mu_{p^\infty}(k(\mu_p))|L_k(\chi^{-1}, 0))| & \text{if } \chi = \omega. \end{cases}$$

Proof. Claim (i) is an immediate consequence of Corollary 8.2 and a remark before this corollary. We shall now prove claim (ii).

When $\chi \neq \omega$, we take a finite place v such that v is prime to p and $\mathrm{N}v \not\equiv \chi(\mathrm{Fr}_v) \pmod p$. We put $T = \{v\}$. Then $(A^T)^\chi = A^\chi$ and $\mathrm{ord}_p L^T_k(\chi^{-1}, 0) = \mathrm{ord}_p L_k(\chi^{-1}, 0)$. Therefore, claim (i) implies the equality in claim (ii).

When $\chi = \omega$, using Chebotarev density theorem we take a finite place v such that v splits completely in $k_\chi = k_\omega = k(\mu_p)$ and $\mathrm{ord}_p|\mu_{p^\infty}(k_\chi)| = \mathrm{ord}_p(\mathrm{N}v - 1)$. We take $T = \{v\}$, then we also have $(A^T)^\chi = A^\chi$ from the exact sequence

$$\mu_{p^\infty}(k_\chi) \longrightarrow (\bigoplus_{w|v} \kappa(w)^\times \otimes \mathbb{Z}_p)^\chi \longrightarrow (A^T)^\chi \longrightarrow A^\chi \longrightarrow 0$$

where w runs over all places of k_χ above v. Therefore, claim (ii) follows from claim (i) in this case, too. □

By combining the argument of Corollary 8.3 with the result of Theorem 1.16 one also directly obtains the following result.

Corollary 8.4. *Assume that at most one p-adic place $\mathfrak{p}$ of k satisfies $\chi(\mathfrak{p}) = 1$. Then the same conclusion as Corollary 8.3 holds.*

Remark 8.5. We note that the formula on A^χ in Corollary 8.3 has not yet been proved in general even in such a semi-simple case (namely the case that the order of χ is prime to p). If no p-adic place $\mathfrak{p}$ satisfies $\chi(\mathfrak{p}) = 1$, this is an immediate consequence of the main conjecture proved by Wiles [54]. Corollary 8.4 shows that this holds even if the set $\{\mathfrak{p} : p\text{-adic place of } k \text{ with } \chi(\mathfrak{p}) = 1\}$ has cardinality one.

8.3. The structure of the class group of a CM field. Now we consider a general CM-field K over a totally real number field k (in particular, we do not assume that $K = k_\chi$).

We assume the condition $(*)$ stated just prior to Theorem 8.1.

We fix a strictly positive integer N. Suppose that v is a place of k such that v is prime to p, v splits completely in K and there is a cyclic extension $F(v)/k$ of degree p^N, which is unramified outside v and in which v is totally ramified. (Note that $F(v)$ is not unique.) We denote by $\mathcal{S}(K)$ the set of such places v and recall that $\mathcal{S}(K)$ is infinite (see [30, Lem. 3.1]).

Suppose now that $V = \{v_1, \dots, v_t\}$ is a subset of $\mathcal{S}(K)$ consisting of t distinct places. We take a cyclic extension $F(v_j)/k$ as above, and put $F = F(v_1)\cdots F(v_t)$ the compositum of fields $F(v_j)$. In particular, F is totally real. We denote by $\mathcal{F}_{t,N}$ the set of all fields F constructed in this way. When $t = 0$, we define $\mathcal{F}_{0,N} = \{k\}$.

We set

$$H := \mathrm{Gal}(KF/K) \cong \mathrm{Gal}(F/k) \cong \prod_{j=1}^{t} \mathrm{Gal}(F(v_j)/k),$$

where the first (restriction) isomorphism is due to the fact that $K \cap F = k$ and the second to the fact that each extension $F(v_j)/k$ is totally ramified at v_j and unramified at all other places.

We fix a generator σ_j of $\mathrm{Gal}(F(v_j)/k)$ and set $S_j := \sigma_j - 1 \in \mathbb{Z}[\mathrm{Gal}(KF/k)]$. Noting that $\mathrm{Gal}(KF/k) = G \times H$ where $G = \mathrm{Gal}(K/k)$, for each element x of $\mathbb{Z}[\mathrm{Gal}(KF/k)] = \mathbb{Z}[G][H]$ we write $x = \sum x_{n_1,\dots,n_t} S_1^{n_1} \cdots S_t^{n_t}$ where each $x_{n_1,\dots,n_t}$ belongs to $\mathbb{Z}[G]$. We then define a map

$$\varphi_V : \mathbb{Z}[\mathrm{Gal}(KF/k)] \to \mathbb{Z}/p^N[G]$$

by sending x to $x_{1,\dots,1}$ modulo p^N and we note that this map is a well-defined homomorphism of G-modules.

We consider $\theta_{KF/k,S\cup V,T}(0) \in \mathbb{Z}[\mathrm{Gal}(KF/k)]$. We define $\Theta^i_{N,S,T}(K/k)$ to be the ideal of $\mathbb{Z}/p^N[G]$ generated by all $\varphi_V(\theta_{KF/k,S\cup V,T}(0)) \in \mathbb{Z}/p^N[G]$ where F

runs over $\mathcal{F}_{t,N}$ such that $t \leq i$. We note that we can compute $\theta_{KF/k,S\cup V,T}(0)$, and hence also $\varphi_V(\theta_{KF/k,S\cup V,T}(0))$, numerically. Taking $F = k$, we know that $\theta_{K/k,S,T}(0) \bmod p^N$ is in $\Theta^i_{N,S,T}(K/k)$ for any $i \geq 0$.
We set $\mathcal{F}_N := \bigcup_{t\geq 0} \mathcal{F}_{t,N}$.
For any abelian extension M/k, if $S = S_\infty(k) \cup S_{\mathrm{ram}}(M/k)$ and T is the empty set, we write $\theta_{M/k}(0)$ for $\theta_{M/k,S,T}(0)$.
We take a character χ of Δ_K such that $\chi \neq \omega$, at first. We take $S = S_\infty(k) \cup S_{\mathrm{ram}}(K/k)$ and $T = \emptyset$. In this case, we know that the χ-component $\theta_{KF/k}(0)^\chi$ is integral, namely is in $O_\chi[\Gamma_K \times H]$. We simply denote the χ-component $\Theta^i_{N,S,\emptyset}(K/k)^\chi$ by $\Theta^i_N(K/k)^\chi$ ($\subset O_\chi[\Gamma_K]$). This ideal $\Theta^i_N(K/k)^\chi$ coincides with the higher Stickelberger ideal $\Theta^{(\delta,N),\chi}_{i,K}$ defined in [30, §8.1].
When $\chi = \omega$, we assume that $K = k(\mu_{p^m})$ for some $m \geq 1$. By using the Chebotarev density theorem we can choose a place v which satisfies all of the following conditions

(i) v splits completely in $k(\mu_p)/k$,
(ii) each place above v of $k(\mu_p)$ is inert in $K/k(\mu_p)$, and
(iii) each place w of K above v satisfies $\mathrm{ord}_p|\mu_{p^\infty}(K)| = \mathrm{ord}_p(\mathrm{N}w - 1)$.

We set $T := \{v\}$. We consider the ω-component $\Theta^i_{N,S,\{v\}}(K/k)^\omega$, which we denoted by $\Theta^i_{N,\{v\}}(K/k)^\omega$

Theorem 8.6. *Let K/k be a finite abelian extension, K a CM-field, and k totally real. Suppose that χ is an odd faithful character of Δ_K, and consider the χ-component of the class group $A(K)^\chi$ which is an $\mathcal{O}_\chi[\Gamma_K]$-module. We assume the condition $(*)$ stated just prior to Theorem 8.1 and the validity of the χ-component of* LTC(FK/k) *for every field F in $\mathcal{F}_N$.*

(i) *Suppose that $\chi \neq \omega$. For any integer $i \geq 0$, we have*
$$\mathrm{Fitt}^i_{\mathcal{O}_\chi[\Gamma_K]/p^N}(A(K)^\chi \otimes \mathbb{Z}/p^N) = \Theta^i_N(K/k)^\chi.$$
(ii) *We assume that $K = k(\mu_{p^m})$ for some $m \geq 1$. For $\chi = \omega$, using a place v as above, we have*
$$\mathrm{Fitt}^i_{\mathcal{O}_\omega[\Gamma_K]/p^N}(A(K)^\omega \otimes \mathbb{Z}/p^N) = \Theta^i_{N,\{v\}}(K/k)^\omega$$
for any $i \geq 0$.

Proof. We first prove claim (i). Since the image of $\theta_{KF/k}(0)$ in $\mathbb{Z}[G]$ is a multiple of $\theta_{K/k}(0)$, $\Theta^0_N(K/k)^\chi$ is a principal ideal generated by $\theta_{K/k}(0)^\chi$. Therefore, this theorem for $i = 0$ follows from Theorem 8.1.
Now suppose that $i > 0$. For a place $v \in \mathcal{S}(K)$, we take a place w of K above v. Put $H(v) = \mathrm{Gal}(F(v)K/K) = \mathrm{Gal}(F(v)/k) \simeq \mathbb{Z}/p^N$. We take a generator σ_v of $H(v)$ and fix it. We define ϕ_v by
$$\begin{aligned}\phi_v : K^\times &\xrightarrow{\mathrm{Rec}_v} (I(H(v))\mathbb{Z}[\mathrm{Gal}(F(v)K/k)]/I(H(v))^2\mathbb{Z}[\mathrm{Gal}(F(v)K/k)]) \\ &= \mathbb{Z}[G] \otimes_{\mathbb{Z}} I(H(v))/I(H(v))^2 \simeq \mathbb{Z}/p^N[G]\end{aligned}$$

Here, the last isomorphism is defined by $\sigma_v - 1 \mapsto 1$, and Rec_v is defined by
$$\mathrm{Rec}_v(a) = \sum_{\tau \in G} \tau^{-1}(\mathrm{rec}_w(\tau a) - 1)$$
as in §5.3 by using the reciprocity map $\mathrm{rec}_w : K_w^\times \to H(v)$ at w. Taking the χ-component of ϕ_v, we obtain
$$\phi_v : (K^\times \otimes \mathbb{Z}/p^N)^\chi \longrightarrow \mathcal{O}_\chi[\Gamma_K]/p^N,$$
which we also denote by ϕ_v.

We take $S = S_\infty(k) \cup S_{\mathrm{ram}}(K/k)$, $T = \emptyset$, and $V = \{v_1, \ldots, v_i\} \in \mathcal{V}_i$. Suppose that $\Phi = \varphi_1 \wedge \cdots \wedge \varphi_i$ where
$$\varphi_j \in \mathcal{H}_\chi = \mathrm{Hom}_{\mathcal{O}_\chi[\Gamma_K]}((\mathcal{O}_{K,S\cup V}^\times \otimes \mathbb{Z}_p)^\chi, \mathcal{O}_\chi[\Gamma_K]/p^N)$$
for $j = 1, \ldots, i$. We take a place w_j of K above v_j for j such that $1 \leq j \leq i$. We denote by $[w_j]$ the class of w_j in $A(K)^\chi$.

By [30, Lem. 10.1], for each integer $j = 1,...,i$ we can choose a place $v'_j \in \mathcal{S}(K)$ that satisfies all of the following conditions;

(a) $[w'_j] = [w_j]$ in $A(K)^\chi$ where w'_j is a place of K above v'_j,

(b) $\varphi_j(x) = \phi_{v'_j}(x)$ for any $x \in (\mathcal{O}_{K,S\cup V}^\times \otimes \mathbb{Z}/p^N)^\chi$,

Here, we used the fact that the natural map $(\mathcal{O}_{K,S\cup V}^\times \otimes \mathbb{Z}/p^N)^\chi \to (K^\times \otimes \mathbb{Z}/p^N)^\chi$ is injective.

Set $V' = \{v'_1, \ldots, v'_i\}$. By property (b), we have
$$\Phi(\epsilon^{V,\chi}_{K/k,S\cup V,\emptyset}) = (\phi_{v'_1} \wedge \cdots \wedge \phi_{v'_i})(\epsilon^{V,\chi}_{K/k,S\cup V,\emptyset}).$$
By property (a), there exists an x_j in $\mathcal{O}_{K,S\cup V\cup V'}^\times$ whose prime decomposition is $(x_j) = w_j(w'_j)^{-1}$ for any j such that $1 \leq j \leq i$. Put $V_{i-1} = \{v_1, ..., v_{i-1}\}$ and $V'_i = \{v_1, ..., v_{i-1}, v'_i\}$. Then
$$\epsilon^{V,\chi}_{K/k,S\cup V,\emptyset} = \epsilon^{V'_i,\chi}_{K/k,S\cup V'_i,\emptyset} + \epsilon^{V_{i-1},\chi}_{K/k,S\cup V_{i-1},\emptyset} \wedge x_i$$
and by using this kind of equation recursively, one deduces that $\epsilon^{V,\chi}_{K/k,S\cup V,\emptyset} - \epsilon^{V',\chi}_{K/k,S\cup V',\emptyset}$ is a sum of elements of the form $\epsilon^{W,\chi}_{K/k,S\cup W,\emptyset} \wedge b_j$ with $|W| = i-1$. Now, by induction on i, we know $\Psi(\epsilon^{W,\chi}_{K/k,S\cup W,\emptyset})$ is in $\Theta_N^{i-1}(K/k)^\chi$ for any Ψ in $\mathrm{Hom}_{\mathcal{O}_\chi[\Gamma_K]}((\mathcal{O}_{K,S\cup W}^\times \otimes \mathbb{Z}_p)^\chi, \mathcal{O}_\chi[\Gamma_K]/p^N)$. Therefore we have
$$(\phi_{v'_1} \wedge \cdots \wedge \phi_{v'_i})(\epsilon^{V,\chi}_{K/k,S\cup V,\emptyset}) \equiv (\phi_{v'_1} \wedge \cdots \wedge \phi_{v'_i})(\epsilon^{V',\chi}_{K/k,S\cup V',\emptyset}) \pmod{\Theta_N^{i-1}(K/k)^\chi}.$$
Set $F = F(v'_1) \cdots F(v'_i)$ and $H = \mathrm{Gal}(FK/K) = \mathrm{Gal}(F/k)$. Then as in §5.3 we can define $\mathrm{Rec}_{V'}(\epsilon^{V',\chi}_{K/k,S\cup V',\emptyset}) \in \mathbb{Z}[G] \otimes (J_{V'})_H$. Let $\varphi_{V'} : \mathbb{Z}[G \times H] \to \mathbb{Z}/p^N[G]$ be the homomorphism defined before Theorem 8.6 by using the generators $\sigma_{v'_i}$ we fixed. This $\varphi_{V'}$ induces a homomorphism
$$\mathbb{Z}[G \times H]/I(H)^{i+1}\mathbb{Z}[G \times H] = \mathbb{Z}[G] \otimes \mathbb{Z}[H]/I(H)^{i+1} \to \mathbb{Z}/p^N[G]$$
and we also denote the composite homomorphism
$$\mathbb{Z}[G] \otimes (J_{V'})_H \to \mathbb{Z}[G] \otimes \mathbb{Z}[H]/I(H)^{i+1} \xrightarrow{\varphi_{V'}} \mathbb{Z}/p^N[G]$$

by $\varphi_{V'}$.

Then by the definitions of these homomorphisms, we have

$$(\phi_{v'_1} \wedge \cdots \wedge \phi_{v'_i})(\epsilon^{V',\chi}_{K/k,S\cup V',\emptyset}) = \varphi_{V'}(\mathrm{Rec}_{V'}(\epsilon^{V',\chi}_{K/k,S\cup V',\emptyset})).$$

By Conjecture 5.4 which is a theorem under our assumptions (Theorem 5.16), we get

$$\varphi_{V'}(\mathrm{Rec}_{V'}(\epsilon^{V',\chi}_{K/k,S\cup V',\emptyset})) = \varphi_{V'}(\theta_{KF/k}(0)^\chi).$$

Combining the above equations, we get

$$\Phi(\epsilon^{V,\chi}_{K/k,S\cup V,\emptyset}) \equiv \varphi_{V'}(\theta_{KF/k}(0)^\chi) \pmod{\Theta_N^{i-1}(K/k)^\chi}.$$

Since $\varphi_{V'}(\theta_{KF/k}(0)^\chi)$, $\Theta_N^{i-1}(K/k)^\chi$ are in $\Theta_N^i(K/k)^\chi$, we get $\Phi(\epsilon^{V,\chi}_{K/k,S\cup V,\emptyset}) \in \Theta_N^i(K/k)^\chi$. It follows from Theorem 8.1 that the left hand side of the equation in Theorem 8.6 (i) is in the right hand side.

On the other hand, suppose that F is in $\mathcal{F}_{t,N}$ with $t \leq i$, and that $V = \{v_1, \ldots, v_t\}$ is the set of ramifying place in F/k. As above, by Theorem 5.16 we have

$$\varphi_V(\theta_{KF/k}(0)^\chi) = \varphi_V(\mathrm{Rec}_V(\epsilon^{V,\chi}_{K/k,S\cup V,\emptyset})) = (\phi_{v_1} \wedge \cdots \wedge \phi_{v_t})(\epsilon^{V,\chi}_{K/k,S\cup V,\emptyset}).$$

Therefore, by Theorem 8.1 we have

$$\varphi_V(\theta_{KF/k}(0)^\chi) \in \mathrm{Fitt}^t_{\mathcal{O}_\chi[\Gamma_K]/p^N}(A(K)^\chi \otimes \mathbb{Z}/p^N).$$

Since $\mathrm{Fitt}^t_{\mathcal{O}_\chi[\Gamma_K]/p^N}(A(K)^\chi \otimes \mathbb{Z}/p^N) \subset \mathrm{Fitt}^i_{\mathcal{O}_\chi[\Gamma_K]/p^N}(A(K)^\chi \otimes \mathbb{Z}/p^N)$, we get

$$\varphi_V(\theta_{KF/k}(0)^\chi) \in \mathrm{Fitt}^i_{\mathcal{O}_\chi[\Gamma_K]/p^N}(A(K)^\chi \otimes \mathbb{Z}/p^N).$$

Thus, the right hand side of the equation in Theorem 8.6 (i) is in the left hand side.

We can prove claim (ii) by the same method. The condition on v is used to show the injectivity of the natural homomorphism $(\mathcal{O}^\times_{K,S\cup V,T} \otimes \mathbb{Z}/p^N)^\omega \to (K^\times \otimes \mathbb{Z}/p^N)^\omega$ with $T = \{v\}$. □

Corollary 8.7. *Let K/k and χ be as in Theorem 8.6. We assume the condition $(*)$ stated just prior to Theorem 8.1 and that there is at most one place $\mathfrak{p}$ of k above p such that $\chi(\mathfrak{p}) = 1$. Then the same conclusion as in Theorem 8.6 holds.*

Proof. It suffices to note that, under the stated conditions, Theorem 1.16 implies that the χ-component of $\mathrm{LTC}(FK/k)$ is valid. □

To give an example of Corollary 8.7 we suppose that K is the m-th layer of the cyclotomic $\mathbb{Z}_p$-extension of $K_{(\Delta)}$ for some strictly positive integer m, and assume that $\chi(\mathfrak{p}) \neq 1$ for any $\mathfrak{p} \mid p$.

Then this assumption implies that the condition $(*)$ is satisfied and so all of the assumptions in Corollary 8.7 are satisfied. Therefore, by taking the projective limit of the conclusion, Corollary 8.7 implies the result of the second author in [30, Th. 2.1].

In this sense, Corollary 8.7 is a natural generalization of the main result in [30].

To state our final result we now set

$$\Theta^i(K/k)^\chi = \varprojlim_N \Theta^i_N(K/k)^\chi \subseteq \mathcal{O}_\chi[\Gamma_K].$$

Then Theorem 8.6 implies that $\mathrm{Fitt}^i_{\mathcal{O}_\chi[\Gamma]}(A(K)^\chi) = \Theta^i(K/k)^\chi$.

Let k_χ be the field corresponding to the kernel of χ as in Corollary 8.3. We denote $\Theta^i(k_\chi/k)^\chi$ by $\Theta^{i,\chi}$, which is an ideal of $\mathcal{O}_\chi$. For $\chi = \omega$, we denote $\varprojlim_N \Theta^i_{N,\{v\}}(k_\chi/k)^\omega$ by $\Theta^{i,\omega}$.

Then Corollary 8.7 implies the following result, which is a generalization of the main result of the second author in [28].

Corollary 8.8. *Set $A^\chi := (\mathrm{Cl}(k_\chi) \otimes \mathbb{Z}_p)^\chi$ as in Corollary 8.3. Assume that there is at most one p-adic place $\mathfrak{p}$ of k such that $\chi(\mathfrak{p}) = 1$ and that the p-adic Iwasawa μ-invariant of K vanishes.*

Then there is an isomorphism of $\mathcal{O}_\chi$-modules of the form $A^\chi \simeq \bigoplus_{i\geq 1} \Theta^{i,\chi}/\Theta^{i-1,\chi}$.

Acknowledgements It is a pleasure for the first author to thank Dick Gross for much encouragement at an early stage of this general project and, in addition, to thank Cornelius Greither for stimulating discussions. The second author would like to thank Cornelius Greither for discussions with him on various topics related to the subjects in this paper. He also thanks Kazuya Kato and Karl Rubin for their beautiful ideas to define several zeta elements, and for stimulating conversations and discussions with them. The third author would like to thank Kazuya Kato for his interest in the works of the third author and for encouragement.

References

[1] S. Bloch, K. Kato, L-functions and Tamagawa numbers of motives, in The Grothendieck Festschrift Vol I, Progress in Math. Vol 86, Birkhäuser (1990), 333-400.

[2] D. Burns, On the values of equivariant Zeta functions of curves over finite fields, Doc. Math., J. DMV 9 (2004) 357-399.

[3] D. Burns, Congruences between derivatives of abelian L-functions at $s = 0$, Invent. math. **169** (2007) 451-499.

[4] D. Burns, Perfecting the nearly perfect, Pure and Applied Math. Q. 4 (2008) 1041-1058 (Jean-Pierre Serre Special Issue, Part I).

[5] D. Burns, Congruences between derivatives of geometric L-functions (with an appendix by D. Burns, K. F. Lai and K-S. Tan), Invent. math. 184 (2011) 221-256.

[6] D. Burns, M. Flach, On Galois structure invariants associated to Tate motives, Amer. J. Math. 120 (1998) 1343-1397.

[7] D. Burns, M. Flach, Tamagawa numbers for motives with (non-commutative) coefficients, Doc. Math. 6 (2001) 501-570.

[8] D. Burns, C. Greither, On the Equivariant Tamagawa Number Conjecture for Tate motives, Invent. math. **153** (2003) 303-359.

[9] D. Burns, M. Kurihara, T. Sano, Iwasawa theory and zeta elements for $\mathbb{G}_m$, preprint (2015) arXiv:1506.07935v3.

[10] P. Cornacchia, C. Greither, Fitting ideals of class groups of real fields with prime power conductor, J. Number Theory 73 (1998) 459-471.

[11] C. W. Curtis, I. Reiner, Methods of Representation Theory, Vol. I, John Wiley and Sons, New York, 1987.

[12] H. Darmon, Thaine's method for circular units and a conjecture of Gross, Canadian J. Math., 47 (1995) 302-317.

[13] C. Emmons, C. Popescu, Special Values of Abelian L-functions at $s = 0$, J. Number Theory 129 (2009) 1350-1365.

[14] M. Flach, On the cyclotomic main conjecture for the prime 2, J. reine angew. Math. **661** (2011) 1-36.

[15] C. Greither, Determining Fitting ideals of minus class groups via the equivariant Tamagawa number conjecture, Compositio Math. 143 (2007) 1399-1426.

[16] C. Greither, R. Kučera, On a conjecture concerning minus parts in the style of Gross, Acta Arith. 132 (2008) 1-48.

[17] C. Greither, R. Kučera, The Minus Conjecture revisited, J. reine angew. Math. 632 (2009) 127-142.

[18] C. Greither, M. Kurihara, Stickelberger elements, Fitting ideals of class groups of CM fields, and dualisation, Math. Zeitschrift 260 (2008), 905-930.

[19] C. Greither, C. Popescu, An Equivariant Main Conjecture in Iwasawa Theory and Applications, J. Algebraic Geom. 24 (2015) 629-692.

[20] B. H. Gross, On p-adic L-series at $s = 0$, J. Fac. Sci. Univ. Tokyo, Sect. IA, 28 (1982) 979-994.

[21] B. H. Gross, On the value of abelian L-functions at $s = 0$, J. Fac. Sci. Univ. Tokyo, Sect. IA, 35 (1988) 177-197.

[22] A. Hayward, A class number formula for higher derivatives of abelian L-functions, Compositio Math. 140 (2004) 99-129.

[23] U. Jannsen, Iwasawa modules up to isomorphism, Advanced Studies in Pure Mathematics 17 (1989) 171-207.

[24] K. Kato, Iwasawa theory and p-adic Hodge theory, Kodai Math. J. 16 no 1 (1993) 1-31.

[25] K. Kato, Lectures on the approach to Iwasawa theory of Hasse-Weil L-functions via B_{dR}, Part I, In: Arithmetical Algebraic Geometry (ed. E. Ballico), Lecture Notes in Math. 1553 (1993) 50-163, Springer, New York, 1993.

[26] S. Y. Kim, On the Equivariant Tamagawa Number Conjecture for Quaternion Fields, Ph. D. thesis, King's College London, 2003.

[27] F. Knudsen, D. Mumford, The projectivity of the moduli space of stable curves I: Preliminaries on 'det' and 'Div', Math. Scand. 39 (1976) 19-55.

[28] M. Kurihara, On the structure of ideal class groups of CM-fields, Documenta Mathematica, Extra Volume Kato (2003), 539-563.

[29] M. Kurihara, On stronger versions of Brumer's conjecture, Tokyo Journal of Math. 34 (2011) 407-428.

[30] M. Kurihara, Refined Iwasawa theory and Kolyvagin systems of Gauss sum type, Proceedings of the London Mathematical Society 104 (2012), 728-769.

[31] M. Kurihara, T. Miura, Ideal class groups of CM-fields with non-cyclic Galois action, Tokyo Journal of Math. 35 (2012) 411-439.

[32] T. Y. Lam, A First Course in Noncommutative Rings (Second Edn.), Grad. Text Math. 131, Springer-Verlag 2001.

[33] S. Lichtenbaum, The Weil-étale topology on schemes over finite fields, Compositio Math. 141 (2005) 689-702.

[34] S. Lichtenbaum, The Weil-étale topology for number rings, Ann. Math. 170 (2009), 657-683.

[35] B. Mazur, K. Rubin, Kolyvagin systems, Mem. Amer. Math. Soc. 168 Number 799 (2004).

[36] B. Mazur, K. Rubin, Refined class number formulas and Kolyvagin systems, Compos. Math. **147** (2011) 56-74.

[37] B. Mazur, K. Rubin, Refined class number formulas for $\mathbb{G}_m$, preprint (2013) arXiv:1312.4053v1.

[38] B. Mazur, J. Tate, Refined Conjectures of the Birch and Swinnerton-Dyer Type, Duke Math. J. 54 (1987) 711-750.

[39] J.S. Milne, Arithmetic Duality Theorems, Perspectives in Mathematics 1, Academic Press, 1986.

[40] D. G. Northcott, Finite free resolutions, Cambridge Univ. Press, Cambridge New York 1976.

[41] T. Ohshita, On the higher Fitting ideals of Iwasawa modules of ideal class groups over real abelian fields, J. Number Theory 135 (2014) 67-138.

[42] C. D. Popescu, Integral and p-adic refinements of the Abelian Stark conjecture, In: Arithmetic of L-functions by C. Popescu, K. Rubin, and A. Silverberg, IAS/Park City Mathematics Series 18, AMS (2011) 45-101.

[43] K. Rubin, Global units and ideal class groups, Invent. math. 89 (1987) 511-526.

[44] K. Rubin, Kolyvagin's system of Gauss sums, In: Arithmetic Algebraic Geometry, G. van der Geer et al eds, Progress in Math 89 (1991) 309-324.

[45] K. Rubin, A Stark Conjecture 'over $\mathbb{Z}$' for abelian L-functions with multiple zeros, Ann. Inst. Fourier 46 (1996) 33-62.

[46] T. Sano, Refined abelian Stark conjectures and the equivariant leading term conjecture of Burns, Compositio Math. 150 (2014) 1809-1835.

[47] D. Solomon, On a construction of p-units in abelian fields, Invent. Math. 109 (1992) 329-350.

[48] K.-S. Tan, Generalized Stark formulae over function fields, Trans. Am. Math. Soc. 361 (2009) 2277-2304.

[49] J. Tate, The cohomology groups of tori in finite Galois extensions of number fields, Nagoya Math. J., 27 (1966) 709-719.
[50] J. Tate, Les Conjectures de Stark sur les Fonctions L d'Artin en $s = 0$ (notes par D. Bernardi et N. Schappacher), Progress in Math., 47, Birkhäuser, Boston, 1984.
[51] J. Tate, Refining Gross's conjecture on the values of abelian L-functions, Contemp. Math. 358 (2004) 189-192.
[52] D. Vallières, The equivariant Tamagawa number conjecture and the extended abelian Stark conjecture, to appear in J. reine u. angew. Math.
[53] J. L. Verdier, Des Catégories Dérivées des Catégories Abéliennes, Astérisque 239, Soc. Math. France, 1996.
[54] A. Wiles, The Iwasawa conjecture for totally real fields, Ann. Math. 131 (1990) 493-540.

David Burns
King's College London
Department of Mathematics
London WC2R 2LS, U.K.
david.burns@kcl.ac.uk

Masato Kurihara
Keio University
Department of Mathematics
3-14-1 Hiyoshi
Kohoku-ku, Yokohama, 223-8522
Japan
kurihara@math.keio.ac.jp

Takamichi Sano
Osaka University
Department of Mathematics
1-1 Machikaneyama-cho
Toyonaka, Osaka, 560-0043
Japan
tkmc310@gmail.com

The Harmonic Analysis of Lattice Counting on Real Spherical Spaces

Bernhard Krötz[1], Eitan Sayag[2], Henrik Schlichtkrull

Received: August 28, 2014
Revised: December 17, 2015

Communicated by Patrick Delorme

Abstract. By the collective name of *lattice counting* we refer to a setup introduced in [10] that aims to establish a relationship between arithmetic and randomness in the context of affine symmetric spaces. In this paper we extend the geometric setup from symmetric to real spherical spaces and continue to develop the approach with harmonic analysis which was initiated in [10].

2010 Mathematics Subject Classification: 22E40, 22E46, 43A85
Keywords and Phrases: Lattice counting, homogeneous spaces, real spherical spaces, spectral analysis

1 Introduction

1.1 Lattice counting

Let us recall from Duke, Rudnick and Sarnak [10] the setup of lattice counting on a homogeneous space $Z = G/H$. Here G is an algebraic real reductive group and $H < G$ an algebraic subgroup such that Z carries an invariant measure. Further we are given a lattice $\Gamma < G$ such that its trace $\Gamma_H := \Gamma \cap H$ in H is a lattice in H.
Attached to invariant measures dh and dg on H and G we obtain an invariant measure $d(gH)$ on Z via Weil-integration:

$$\int_Z \Big(\int_H f(gh)dh\Big)\ d(gH) = \int_G f(g)\ dg \qquad (f \in C_c(G))\,.$$

[1]Supported by ERC Advanced Investigators Grant HARG 268105
[2]Partially supported by ISF 1138/10 and ERC 291612

Likewise the measures dg and dh give invariant measures $d(g\Gamma)$ and $d(h\Gamma_H)$ on $Y := G/\Gamma$ and $Y_H := H/\Gamma_H$. We pin down the measures dg and dh and hence $d(gH)$ by the request that Y and Y_H have volume one.
Further we are given a family $\mathcal{B}$ of "balls" $B_R \subset Z$ depending on a parameter $R \geq 0$. At this point we are rather imprecise about the structure of these balls and content us with the property that they constitute an exhausting family of compact sets as $R \to \infty$.
Let $z_0 = H \in Z$ be the standard base point. The *lattice counting problem* for $\mathcal{B}$ consists of the determination of the asymptotic behavior of the density of $\Gamma \cdot z_0$ in balls $B_R \subset Z$, as the radius $R \to \infty$. By *main term counting* for $\mathcal{B}$ we understand the statement that the asymptotic density is 1. More precisely, with

$$N_R(\Gamma, Z) := \#\{\gamma \in \Gamma/\Gamma_H \mid \gamma \cdot z_0 \in B_R\}$$

and $|B_R| := \mathrm{vol}_Z(B_R)$ we say that main term counting holds if

$$N_R(\Gamma, Z) \sim |B_R| \quad (R \to \infty). \tag{1.1}$$

1.2 Relevant previous works

The main term counting was established in [10] for symmetric spaces G/H and certain families of balls, for lattices with Y_H compact. Furthermore, the main term counting in the case where Y_H is non-compact was proven using a hypothesis on regularization of periods of Eisenstein series, whose proof remains unpublished. In subsequent work Eskin and McMullen [11] removed the obstruction that Y_H is compact and presented an ergodic approach. Later Eskin, Mozes and Shah [12] refined the ergodic methods and discovered that main term counting holds for a wider class of reductive spaces: For reductive algebraic groups G, H defined over $\mathbb{Q}$ and arithmetic lattices $\Gamma < G(\mathbb{Q})$ it is enough to request that the identity component of H is not contained in a proper parabolic subgroup of G which is defined over $\mathbb{Q}$ and that the balls B_R satisfy a certain condition of *non-focusing*.
In these works the balls B_R are constructed as follows. All spaces considered are affine in the sense that there exists a G-equivariant embedding of Z into the representation module V of a rational representation of G. For any such embedding and any norm on the vector space V, one then obtains a family of balls B_R on Z by intersection with the metric balls in V. For symmetric spaces all families of balls produced this way are suitable for the lattice counting, but in general one needs to assume non-focusing in addition. In particular all maximal reductive subgroups satisfy all the conditions and hence fulfill the main term counting.

1.3 Real spherical spaces

In this paper we investigate the lattice counting for a real spherical space Z, that is, it is requested that the action of a minimal parabolic subgroups $P < G$ on

Z admits an open orbit. In addition we assume that H is reductive and remark that with our standing assumption that Z is unimodular this is automatically satisfied for a spherical space when the Lie algebra $\mathfrak{h}$ of H is self-normalizing (see [17], Cor. 9.10).
Our approach is based on spectral theory and is a natural continuation to [10]. We consider a particular type of balls which are intrinsically defined by the geometry of Z (and thus not related to a particular representation V as before).

1.3.1 Factorization of spherical spaces

In the spectral approach it is of relevance to get a control over intermediate subgroups $H < H^\star < G$ which arise in the following way: Given a unitary representation $(\pi, \mathcal{H})$ one looks at the smooth vectors $\mathcal{H}^\infty$ and its continuous dual $\mathcal{H}^{-\infty}$, the distribution vectors. The space $(\mathcal{H}^{-\infty})^H$ of H-invariant distribution vectors is of fundamental importance. For all pairs $(v, \eta) \in \mathcal{H}^\infty \times (\mathcal{H}^{-\infty})^H$ one obtains a smooth function on Z, a *generalized matrix-coefficient*, via

$$m_{v,\eta}(z) = \eta(g^{-1} \cdot v) \qquad (z = gH \in Z)\,. \tag{1.2}$$

The functions (1.2) are the building blocks for the harmonic analysis on Z. The stabilizer H_η in G of $\eta \in (\mathcal{H}^{-\infty})^H$ is a closed subgroup which contains H, but in general it can be larger than H even if π is non-trivial.
Let us call $Z^\star = G/H^\star$ a *factorization* of Z if $H < H^\star$ and $Z^\star$ is unimodular. For a general real spherical space Z the homogeneous spaces $Z_\eta = G/H_\eta$ can happen to be non-unimodular (see [19] for H the Iwasawa N-subgroup). However there is a large subclass of real spherical spaces which behave well under factorization. Let us call a factorization co-compact if $H^\star/H$ is compact and *basic* if (up to connected components) $H^\star$ is of the form $H_I := HI$ for a normal subgroup $I \triangleleft G$. Finally we call a factorization *weakly basic* if it is obtained by a composition of a basic and a co-compact factorization.

1.3.2 Wavefront spherical spaces

A real spherical space is called *wavefront* if the attached compression cone is a quotient of a closed Weyl-chamber. The relevant definitions will be recalled in Section 3. Many real spherical spaces are wavefront: all symmetric spaces and all Gross-Prasad type spaces $G \times H/H$ (see (3.2) - (3.4)) are wavefront.[3]
The terminology *wavefront* originates from [24] because wavefront real spherical spaces satisfy the "wavefront lemma" of Eskin-McMullen (see [11], [18]) which is fundamental in the approach of [11] to lattice counting.
On the geometric side wavefront real spherical spaces enjoy the following property from [19]: All Z_η are unimodular and the factorizations of the type Z_η are precisely the weakly basic factorizations of Z.

[3]Also, if Z is complex, then of the 78 cases in the list of [4], the non-wavefront cases are (11), (24), (25), (27), (39-50), (60), (61)

On the spectral level wavefront real spherical spaces are distinguished by the following integrability property, also from [19]: The generalized matrix coefficients $m_{v,\eta}$ of (1.2) belong to $L^p(Z_\eta)$ for some $1 \leq p < \infty$ only depending on π and η.

1.3.3 Main term counting

In the theorem below we assume that Z is a wavefront real spherical space of reductive type. For simplicity we also assume that all compact normal subgroups of G are finite.
Using soft techniques from harmonic analysis and a general property of decay from [21], our first result (see Section 5) is:

Theorem A. Let $Z = G/H$ be as above, and assume that $Y = G/\Gamma$ is compact. Then main term counting (1.1) holds.

Since wavefront real spherical spaces satisfy the wavefront lemma by [18], Section 6, this theorem could also be derived with the ergodic method of [11]. In the current context the main point is thus the proof by harmonic analysis.
To remove the assumption that Y is compact and to obtain error term bounds for the lattice counting problem we need to apply more sophisticated tools from harmonic analysis. This will be discussed in the next paragraph with some extra assumptions on G/H.

1.4 Error Terms

The problem of determining the error term in counting problems is notoriously difficult and in many cases relies on deep arithmetic information. Sometimes, like in the Gauss circle problem, some error term is easy to establish but getting an optimal error term is a very difficult problem.
We restrict ourselves to the cases where the cycle H/Γ_H is compact.[4] To simplify the exposition here we assume in addition that $\Gamma < G$ is irreducible, i.e. there do not exist non-trivial normal subgroups G_1, G_2 of G and lattices $\Gamma_i < G_i$ such that $\Gamma_1\Gamma_2$ has finite index in Γ.
The error we study is measure theoretic in nature, and will be denoted here as $\mathrm{err}(R,\Gamma)$. Thus, $\mathrm{err}(R,\Gamma)$ measures the deviation of two measures on $Y = \Gamma\backslash G$, the counting measure arising from lattice points in a ball of radius R, and the invariant measure $d\mu_Y$ on Y. More precisely, with $\mathbf{1}_R$ denoting the characteristic function of B_R we consider the densities

$$F_R^\Gamma(g\Gamma) := \frac{\sum_{\gamma\in\Gamma/\Gamma_H} \mathbf{1}_R(g\gamma H)}{|B_R|}.$$

Then,

$$\mathrm{err}(R,\Gamma) = ||F_R^\Gamma - d\mu_Y||_1,$$

[4]After a theory for regularization of H-periods of Eisenstein series is developed, one can drop this assumption.

where $||\cdot||_1$ denotes the total variation of the signed measure. Notice that $|F_R^\Gamma(e\Gamma)-1| = \frac{|N_R(\Gamma,Z)-|B_R||}{|B_R|}$ is essentially the error term for the pointwise count (1.1).

Our results on the error term $\mathrm{err}(R,\Gamma)$ allows us to deduce results toward the error term in the smooth counting problem, a classical problem that studies the quantity

$$\mathrm{err}_{pt,\alpha}(R,\Gamma) = |B_R||F_{\alpha,R}^\Gamma(e\Gamma)-1|$$

where $\alpha \in C_c^\infty(G)$ is a positive smooth function of compact support (with integral one) and $F_{\alpha,R}^\Gamma = \alpha * F_R^\Gamma$. See Remark 7.2 for the comparison of $\mathrm{err}(R,\Gamma)$ with $\mathrm{err}_{pt,\alpha}(R,\Gamma)$.

To formulate our result we introduce the exponent $p_H(\Gamma)$ (see (6.2)), which measures the worst L^p-behavior of any generalized matrix coefficient associated with a spherical unitary representation π, which is H-distinguished and occurs in the automorphic spectrum of $L^2(\Gamma\backslash G)$. We first state our result for the non-symmetric case of triple product spaces, which is Theorem 8.2 from the body of the paper.

Theorem B. Let $Z = G_0^3/\operatorname{diag}(G_0)$ for $G_0 = \mathrm{SO}_e(1,n)$ and assume that H/Γ_H is compact. For all $p > p_H(\Gamma)$ there exists a $C = C(p) > 0$ such that

$$\mathrm{err}(R,\Gamma) \le C|B_R|^{-\frac{1}{(6n+3)p}}$$

for all $R \ge 1$. (In particular, main term counting holds in this case). Furthermore, in regards to smooth counting, for any $\alpha \in C_c^\infty(G)$ and for all $p > p_H(\Gamma)$ there exists a $C = C(p,\alpha) > 0$ such that

$$\mathrm{err}_{pt,\alpha}(R,\Gamma) \le C|B_R|^{1-\frac{1}{(6n+3)p}}$$

for all $R \ge 1$.

To the best of our knowledge this is the first error term obtained for a non-symmetric space. The crux of the proof is locally uniform comparison between L^p and L^∞ norms of generalized matrix coefficients $m_{v,\eta}$ which is achieved by applying the model of [3] and [9] for the triple product functional η in spherical principal series.

It is possible to obtain error term bounds under a certain technical hypothesis introduced in Section 6 and refered to as Hypothesis A. This hypothesis in turn is implied by a conjecture on the analytic structure of families of Harish-Chandra modules which we explain in Section 9.1. The conjecture and hence the hypothesis appear to be true for symmetric spaces but requires quite a technical tour de force. In general, the techniques currently available do not allow for an elegant and efficient solution. Under this hypothesis we show that:

Theorem C. Let Z be wavefront real spherical space for which Hypothesis A is valid. Assume also

- G is semisimple with no compact factors

- Γ is arithmetic and irreducible
- $\Gamma_H = H \cap \Gamma$ is co-compact in H.
- $p > p_H(\Gamma)$
- $k > \frac{rank(G/K)+1}{2} \dim(G/K) + 1$

Then, there exists a constant $C = C(p,k) > 0$ such that

$$\operatorname{err}(R,\Gamma) \leq C|B_R|^{-\frac{1}{(2k+1)p}}$$

for all $R \geq 1$. Moreover, if $Y = \Gamma\backslash G$ is compact one can replace the third condition by $k > \dim(G/K) + 1$.

The existence of a non-quantitative error term for symmetric spaces was established in [1] and improved in [14].
We note that in case of the hyperbolic plane our error term is still far from the quality of the bound of A. Selberg. This is because we only use a weak version of the trace formula, namely Weyl's law, and use simple soft Sobolev bounds between eigenfunctions on Y.

2 Reductive homogeneous spaces

In this section we review a few facts on reductive homogeneous spaces: the Mostow decomposition, the associated geometric balls and their factorizations.
We use the convention that real Lie groups are denoted by upper case Latin letters, e.g A, B, C, and their Lie algebras by the corresponding lower case German letter $\mathfrak{a}$, $\mathfrak{b}$, $\mathfrak{c}$.
Throughout this paper G will denote an algebraic real reductive group and $H < G$ is an algebraic subgroup. We form the homogeneous space $Z = G/H$ and write $z_0 = H$ for the standard base point.
Furthermore, unless otherwise mentioned we assume that H is reductive in G, that is, the adjoint representation of H on $\mathfrak{g}$ is completely reducible. In this case we say that G/H is *of reductive type.*
Let us fix a maximal compact subgroup $K < G$ for which we assume that the associated Cartan involution θ leaves H invariant (see the references to [21], Lemma 2.1). Attached to θ is the infinitesimal Cartan decomposition $\mathfrak{g} = \mathfrak{k} + \mathfrak{s}$ where $\mathfrak{s} = \mathfrak{k}^\perp$ is the orthogonal complement with respect to a non-degenerate invariant bilinear form κ on $\mathfrak{g}$ which is positive definite on $\mathfrak{s}$ (if $\mathfrak{g}$ is semi-simple, then we can take for κ the Cartan-Killing form). Further we set $\mathfrak{q} := \mathfrak{h}^\perp$.

2.1 Mostow decomposition

We recall Mostow's polar decomposition:

$$K \times_{H\cap K} \mathfrak{q} \cap \mathfrak{s} \to Z, \quad [k, X] \mapsto k \exp(X) \cdot z_0 \tag{2.1}$$

which is a homeomorphism. With that we define

$$\|k\exp(X)\cdot z_0\|_Z = \|X\| := \kappa(X,X)^{\frac{1}{2}}$$

for $k \in K$ and $X \in \mathfrak{q}\cap\mathfrak{s}$.

2.2 GEOMETRIC BALLS

The problem of lattice counting in Z leads to a question of exhibiting natural exhausting families of compact subsets. We use balls which are intrinsically defined by the geometry of Z.
We define the *intrinsic ball* of radius $R > 0$ on Z by

$$B_R := \{z \in Z \mid \|z\|_Z < R\}\,.$$

Write B_R^G for the intrinsic ball of $Z = G$, that is, if $g = k\exp(X)$ with $k \in K$ and $X \in \mathfrak{s}$, then we put $\|g\|_G = \|X\|$ and define B_R^G accordingly.
Our first interest is the growth of the volume $|B_R|$ for $R \to \infty$. We have the following upper bound.

LEMMA 2.1. *There exists a constant $c > 0$ such that:*

$$|B_{R+r}| \leq e^{cr}|B_R|$$

for all $R \geq 1, r \geq 0$.

Proof. Recall the integral formula

$$\int_Z f(z)\,dz = \int_K \int_{\mathfrak{q}\cap\mathfrak{s}} f(k\exp(X).z_0)\delta(X)\,dX\,dk, \tag{2.2}$$

for $f \in C_c(Z)$, where $\delta(Y)$ is the Jacobian at (k,Y) of the map (2.1). It is independent of k because dz is invariant. Then

$$|B_R| = \int_{X\in\mathfrak{q}\cap\mathfrak{s},\|X\|<R} \delta(X)\,dX\,.$$

Hence it suffices to prove that there exists $c > 0$ such that

$$\int_0^{R+r} \delta(tX)t^{l-1}\,dt \leq e^{cr}\int_0^R \delta(tX)t^{l-1}\,dt$$

for all $X \in \mathfrak{q}\cap\mathfrak{s}$ with $\|X\| = 1$. Here $l = \dim\mathfrak{q}\cap\mathfrak{s}$. Equivalently, the function

$$R \mapsto e^{-cR}\int_0^R \delta(tX)t^{l-1}\,dt$$

is decreasing, or by differentiation,

$$\delta(RX)R^{l-1} \leq c\int_0^R \delta(tX)t^{l-1}\,dt$$

for all R. The latter inequality is established in [12, Lemma A.3] with c independent of X. $\square$

Further we are interested how the volume behaves under distortion by elements from G.

Lemma 2.2. *For all $r, R > 0$ one has $B_r^G B_R \subset B_{R+r}$.*

To prove the lemma we first record that:

Lemma 2.3. *Let $z = gH \in Z$. Then $\|z\|_Z = \inf_{h\in H} \|gh\|_G$.*

Proof. It suffices to prove that $\|\exp(X)h\|_G \geq \|X\|$ for $X \in \mathfrak{q} \cap \mathfrak{s}$, $h \in H$, and by Cartan decomposition of H, we may assume $h = \exp(T)$ with $T \in \mathfrak{h} \cap \mathfrak{s}$. Thus we have reduced to the statement that

$$\|\exp(X)\exp(T)\|_G \geq \|\exp(X)\|_G$$

for $X \perp T$ in $\mathfrak{s}$. In order to see this, we note that for each $g \in G$ the norm $\|g\|_G$ is the length of the geodesic in $K\backslash G$ which joins the origin x_0 to $x_0 g$. More generally the geodesic between $x_0 g_1$ and $x_0 g_2$ has length $\|g_2 g_1^{-1}\|_G$. Hence $c = \|\exp(X)\exp(T)\|_G$ is the distance from $A = x_0\exp(-T)$ to $B = x_0\exp(X)$. As $X \perp T$ the points A and B form a right triangle with $C = x_0$. The hypotenuse has length c and the leg CB has length $a = \|\exp(X)\|$. As the sectional curvatures are non-positive we have $a^2 + b^2 \leq c^2$. In particular $a \leq c$. □

In particular, it follows that

$$\|gz\|_Z \leq \|z\|_Z + \|g\|_G \quad (z \in Z, g \in G) \tag{2.3}$$

and Lemma 2.2 follows.

Remark 2.4. Observe that the norm $\|\cdot\|_G$ on G depends on the chosen Cartan decomposition θ. However, by applying (2.3) with $Z = G$ one sees that the norm obtained with a conjugate θ' of θ will satisfy

$$\|g\|_G' \leq \|g\|_G + c, \qquad \|g\|_G \leq \|g\|_G' + c' \tag{2.4}$$

for all $g \in G$ with some constants $c, c' \geq 0$.
For the definition of $\|\cdot\|_Z$ we assumed that θ leaves H invariant. If instead we use the identity in Lemma 2.3 as the definition of $\|\cdot\|_Z$ then this assumption can be avoided. In any case, it follows that the norms on Z obtained from two different Cartan involutions will satisfy similar inequalities as (2.4). The corresponding families of balls are then also compatible,

$$B_R \subset B'_{R+c}, \qquad B'_R \subset B_{R+c'},$$

for all $R > 0$.

2.3 Factorization

By a (reductive) factorization of $Z = G/H$ we understand a homogeneous space $Z^\star = G/H^\star$ with $H^\star$ an algebraic subgroup of G such that

- $H^\star$ is reductive.
- $H \subset H^\star$.

A factorization is called *compact* if $Z^\star$ is compact, and *co-compact* if the fiber space $\mathcal{F} := H^\star/H$ is compact. It is called *proper* if $\dim H < \dim H^\star < \dim G$.

Lemma 2.5. *Let $Z = G/H \to Z^\star = G/H^\star$ be a factorization. Then the following assertions are equivalent:*

1. *$Z \to Z^\star$ is co-compact.*
2. *There exist a compact subgroup $K^\star < H^\star$ such that $K^\star H = H^\star$.*
3. *There exists a compact subalgebra $\mathfrak{k}^\star < \mathfrak{h}^\star$ such that $\mathfrak{h}^\star = \mathfrak{k}^\star + \mathfrak{h}$ and $\exp(\mathfrak{k}^\star) < H^\star$ compact.*

Proof. First (1) implies (2) by the Mostow decomposition of the reductive homogeneous space $H^\star/H$. Clearly (2) implies (3) as the multiplication map $K^\star \times H \to H^\star$ needs to be submersive by Sard's theorem. Finally, for (3) implies (1) we observe that $H^\star/H$ has finitely many components and $\exp(\mathfrak{k}^\star)H$ is compact and open in there. □

Let $\mathcal{F} \to Z \to Z^\star$ be a factorization of Z. We write $B_R^\star$ and $\mathcal{B}_R^{\mathcal{F}}$ for the intrinsic balls in $Z^\star$ and $\mathcal{F}$, respectively.

Lemma 2.6. *We have $B_R^\star = B_R H^\star/H^\star$ and $B_R^{\mathcal{F}} = B_R \cap \mathcal{F}$.*

Proof. Follows from Lemma 2.3. □

For a compactly supported bounded measurable function ϕ on Z we define the fiberwise integral

$$\phi^{\mathcal{F}}(gH^\star) := \int_{H^\star/H} \phi(gh^\star)\ d(h^\star H)$$

and recall the integration formula

$$\int_Z \phi(gH)\ d(gH) = \int_{Z^\star} \phi^{\mathcal{F}}(gH^\star)\ d(gH^\star) \tag{2.5}$$

under appropriate normalization of measures. Consider the characteristic function $\mathbf{1}_R$ of B_R and note that its fiber average $\mathbf{1}_R^{\mathcal{F}}$ is supported in the compact ball $B_R^\star$. We say that the family of balls $(B_R)_{R>0}$ *factorizes well to* $Z^\star$ provided for all compact subsets $Q \subset G$

$$\lim_{R\to\infty} \frac{\sup_{g\in Q} \mathbf{1}_R^{\mathcal{F}}(gH^\star)}{|B_R|} = 0\,. \tag{2.6}$$

Observe that for all compact subsets Q there exists an $R_0 = R_0(Q) > 0$ such that

$$\sup_{g\in Q} \mathbf{1}_R^{\mathcal{F}}(gH^\star) \le |B_{R+R_0}^{\mathcal{F}}|$$

by Lemma 2.2. Thus the balls B_R factorize well provided

$$\lim_{R\to\infty} \frac{|B_{R+R_0}^{\mathcal{F}}|}{|B_R|} = 0\,. \tag{2.7}$$

for all $R_0 > 0$.

Remark 2.7. The condition that the balls B_R factorize well is closely related to the non-focusing condition (Definition 1.14 in [12]). Thus, in the case of semi-simple connected H, the non-focusing condition of the intrinsic balls is implied by the condition that they factorize well to all factorizations.

2.4 Basic factorizations

There is a special class of factorizations with which we are dealing with in the sequel. ¿From now on we assume that $\mathfrak{g}$ is semi-simple and write

$$\mathfrak{g} = \mathfrak{g}_1 \oplus \ldots \oplus \mathfrak{g}_m$$

for the decomposition into simple ideals. For a reductive subalgebra $\mathfrak{h} < \mathfrak{g}$ and a subset $I \subset \{1, \ldots, m\}$ we define the reductive subalgebra

$$\mathfrak{h}_I := \mathfrak{h} + \mathfrak{g}_I = \mathfrak{h} + \bigoplus_{i\in I} \mathfrak{g}_i\,. \tag{2.8}$$

We say that the factorization is *basic* provided that $\mathfrak{h}^* = \mathfrak{h}_I$ for some I. Finally we call a factorization *weakly basic* if it is built from consecutive basic and co-compact factorizations, that is, there exists a sequence

$$\mathfrak{h}^\star = \mathfrak{h}^k \supset \cdots \supset \mathfrak{h}^0 = \mathfrak{h} \tag{2.9}$$

of reductive subalgebras such that for each i we have $\mathfrak{h}^i = (\mathfrak{h}^{i-1})_I$ for some I or $\mathfrak{h}^i/\mathfrak{h}^{i-1}$ is compact. The following lemma shows that in fact it suffices with $k \le 2$.

Lemma 2.8. *Let $Z \to Z^\star$ be a weakly basic factorization. Then there exists an intermediate factorization $Z \to Z_b \to Z^\star$ such that $Z \to Z_b$ is basic and $Z_b \to Z^\star$ co-compact.*

Proof. Let a sequence (2.9) of factorizations which are consecutively basic or compact be given. We first observe that two consecutive basic factorizations make up for a single basic factorization, and likewise two consecutive compact factorizations yield a single compact factorization by Lemma 2.5. Hence it suffices to prove that we can modify a string

$$\mathfrak{h}^{i+2} \supset \mathfrak{h}^{i+1} \supset \mathfrak{h}^i$$

with $\mathfrak{h}^{i+2}/\mathfrak{h}^{i+1}$ basic and $\mathfrak{h}^{i+1}/\mathfrak{h}^{i}$ compact to

$$\mathfrak{h}^{i+2} \supset \mathfrak{h}_b^{i+1} \supset \mathfrak{h}^i$$

with $\mathfrak{h}^{i+2}/\mathfrak{h}_b^{i+1}$ compact and $\mathfrak{h}_b^{i+1}/\mathfrak{h}^i$ basic.
We have $\mathfrak{h}^{i+2} = \mathfrak{h}^{i+1} + \mathfrak{g}_I$ for some I, and by Lemma 2.5 that $\mathfrak{h}^{i+1} = \mathfrak{h}^i + \mathfrak{c}$ with $\mathfrak{c}$ compact. Then $\mathfrak{h}_b^{i+1} := \mathfrak{h}^i + \mathfrak{g}_I$ is a reductive subalgebra and a basic factorization of $\mathfrak{h}^i$. Furthermore $\mathfrak{h}^{i+2} = \mathfrak{h}_b^{i+1} + \mathfrak{c}$. This establishes the lemma □

3 Wavefront real spherical spaces

We assume that Z is real spherical, i.e. a minimal parabolic subgroup $P < G$ has an open orbit on Z. It is no loss of generality to assume that $PH \subset G$ is open, or equivalently that $\mathfrak{g} = \mathfrak{h} + \mathfrak{p}$.
If L is a real algebraic group, then we write $L_{\rm n}$ for the normal subgroup of L which is generated by all unipotent element. In case L is reductive we observe that $\mathfrak{l}_{\rm n}$ is the sum of all non-compact simple ideals of $\mathfrak{l}$.
According to [20] there is a unique parabolic subgroup $Q \supset P$ with the following two properties:

- $QH = PH$.
- There is a Levi decomposition $Q = LU$ with $L_{\rm n} \subset Q \cap H \subset L$.

Following [20] we call Q a Z-adapted parabolic subgroup.
Having fixed L we let $L = K_L A_L N_L$ be an Iwasawa decomposition of L. We choose an Iwasawa decomposition $G = KAN$ which inflates the one of L, i.e. $K_L < K, A_L = A$ and $N_L < N$. Further we may assume that N is the unipotent radical of the minimal parabolic P.

Remark 3.1. It should be noted that the assumption on the Cartan decomposition θ, which was demanded in Section 2.2, may be overruled by the above requirement to K. However, it follows from Remark 2.4 that the balls B_R can still be defined, and that the difference does not disturb the lattice counting on Z.

Set $A_H := A \cap H$ and put $A_Z = A/A_H$. We recall that $\dim A_Z$ is an invariant of the real spherical space, called the real rank (see [20]).
In [18], Section 6, we defined the notion of *wavefront* for a real spherical space, which we quickly recall. Attached to Z is a geometric invariant, the so-called compression cone which is a closed and convex subcone $\mathfrak{a}_Z^-$ of $\mathfrak{a}_Z$. It is defined as follows. Write $\Sigma_\mathfrak{u}$ for the space of $\mathfrak{a}$-weights of the $\mathfrak{a}$-module $\mathfrak{u}$ and let $\overline{\mathfrak{u}}$ denote the corresponding sum of root spaces for $-\Sigma_\mathfrak{u}$. According to [20] there exists a linear map

$$T : \oplus_{\alpha\in\Sigma_\mathfrak{u}} \mathfrak{g}^{-\alpha} = \overline{\mathfrak{u}} \to \mathfrak{l}_H^\perp \oplus \mathfrak{u} \subset \oplus_{\beta\in\{0\}\cup\Sigma_\mathfrak{u}} \mathfrak{g}^\beta \qquad (3.1)$$

such that $\mathfrak{h} = \mathfrak{l}\cap\mathfrak{h} + \{\overline{X} + T(\overline{X}) \mid \overline{X} \in \overline{\mathfrak{u}}\}$. Here $\mathfrak{l}_H^\perp$ denotes the orthocomplement of $\mathfrak{l}\cap\mathfrak{h}$ in $\mathfrak{l}$. For each pair α, β we denote by

$$T_{\alpha,\beta} : \mathfrak{g}^{-\alpha} \to \mathfrak{g}^{\beta}$$

the map obtained from T by restriction to $\mathfrak{g}^{-\alpha}$ and projection to $\mathfrak{g}^{\beta}$. Then $T = \sum_{\alpha,\beta} T_{\alpha,\beta}$ and by definition

$$\mathfrak{a}_Z^- = \{X \in \mathfrak{a} \mid (\alpha+\beta)(X) \geq 0,\ \forall \alpha, \beta \text{ with } T_{\alpha,\beta} \neq 0\}.$$

It follows from (3.1) that $\alpha + \beta$ vanishes on $\mathfrak{a}_H$ if $T_{\alpha,\beta} \neq 0$. Hence $\mathfrak{a}_Z^- \subset \mathfrak{a}_Z$. If one denotes by $\mathfrak{a}^- \subset \mathfrak{a}$ the closure of the negative Weyl chamber, then $\mathfrak{a}^- + \mathfrak{a}_H \subset \mathfrak{a}_Z^-$ and by definition Z is wavefront if

$$\mathfrak{a}^- + \mathfrak{a}_H = \mathfrak{a}_Z^-.$$

Let us mention that many real spherical spaces are wavefront; for example all symmetric spaces and all Gross-Prasad type spaces $Z = G \times H/H$ with (G, H) one of the following

$$(\mathrm{GL}_{n+1}(\mathbb{C}), \mathrm{GL}_n(\mathbb{C})),\ (\mathrm{GL}_{n+1}(\mathbb{R}), \mathrm{GL}_n(\mathbb{R})), \tag{3.2}$$

$$(\mathrm{GL}_{n+1}(\mathbb{H}), \mathrm{GL}_n(\mathbb{H})),\ (\mathrm{U}(p+1,q), \mathrm{U}(p,q)), \tag{3.3}$$

$$(\mathrm{SO}(n+1,\mathbb{C}), \mathrm{SO}(n,\mathbb{C})),\ (\mathrm{SO}(p+1,q), \mathrm{SO}(p,q))\,. \tag{3.4}$$

We recall from [18] the polar decomposition for real spherical spaces

$$Z = \Omega A_Z^- F \cdot z_0 \tag{3.5}$$

where

- Ω is a compact set of the type $F'K$ with $F' \subset G$ a finite set.
- $F \subset G$ is a finite set with the property that $F \cdot z_0 = T \cdot z_0 \cap Z$ where $T = \exp(i\mathfrak{a})$ and the intersection is taken in $Z_\mathbb{C} = G_\mathbb{C}/H_\mathbb{C}$.

3.1 Volume growth

Define $\rho_Q \in \mathfrak{a}^*$ by $\rho_Q(X) = \frac{1}{2}\mathrm{tr}(\mathrm{ad}_\mathfrak{u} X)$, $X \in \mathfrak{a}$. It follows from the unimodularity of Z and the local structure theorem that $\rho_Q|_{\mathfrak{a}_H} = 0$, i.e. $\rho_Q \in \mathfrak{a}_Z^* = \mathfrak{a}_H^\perp$.

Lemma 3.2. *Let $Z = G/H$ be a wavefront real spherical space. Then*

$$|B_R| \asymp \sup_{\substack{X\in\mathfrak{a} \\ \|X\|\leq R}} e^{2\rho_Q(X)} = \sup_{\substack{X\in\mathfrak{a}_Z^- \\ \|X\|\leq R}} e^{-2\rho_Q(X)}\,. \tag{3.6}$$

Here the expression $f(R) \asymp g(R)$ signifies that the ratio $\frac{f(R)}{g(R)}$ remains bounded below and above as R tends to infinity.

Proof. First note that the equality in (3.6) is immediate from the wavefront assumption.
Let us first show the lower bound, i.e. there exists a $C > 0$ such that for all $R > 0$ one has

$$|B_R| \geq C \sup_{\substack{X \in \mathfrak{a} \\ \|X\| \leq R}} e^{2\rho_Q(X)}.$$

For that we recall the volume bound from [19], Prop. 4.2: for all compact subsets $B \subset G$ with non-empty interior there exists a constant $C > 0$ such that $\mathrm{vol}_Z(Ba \cdot z_0) \geq Ca^{2\rho_Q}$ for all $a \in A_Z^-$. Together with the polar decomposition (3.5) this gives us the lower bound.
As for the upper bound let

$$\mathfrak{a}_R^- := \{X \in \mathfrak{a}^- \mid \|X\| \leq R\}.$$

Observe that $B_R \subset B'_R := KA_R^- K \cdot z_0$. In the sequel it is convenient to realize A_Z as a subgroup of A (and not as quotient): we identify A_Z with $A_H^\perp \subset A$. The upper bound will follow if we can show that

$$|B'_R| \leq C \sup_{\substack{X \in \mathfrak{a} \\ \|X\| \leq R}} e^{2\rho_Q(X)} \qquad (R > 0).$$

for some constant $C > 0$. This in turn will follow from the argument for the upper bound in the proof of Prop. 4.2 in [19]: in this proof we considered for $a \in A_Z^-$ the map

$$\Phi_a : K \times \Omega_A \times \Xi \to G, \quad (k, b, X) \mapsto kb \exp(\mathrm{Ad}(a)X)$$

where $\Omega_A \subset A$ is a compact neighborhood of $\mathbf{1}$ and $\Xi \subset \mathfrak{h}$ is a compact neighborhood of 0. It was shown that the Jacobian of Φ_a, that is $\sqrt{\det(d\Phi_a d\Phi_a^t)}$, is bounded by $Ca^{-2\rho_Q}$. Now this bounds holds as well for the right K-distorted map

$$\Psi_a : K \times \Omega_A \times K \times \Xi \to G, \quad (k, b, k', X) \mapsto kb \exp(\mathrm{Ad}(ak')X).$$

The reason for that comes from an inspection of the proof; all what is needed is the following fact: let $d := \dim \mathfrak{h}$ and consider the action of $\mathrm{Ad}(a)$ on $V = \bigwedge^d \mathfrak{g}$. Then for $a \in A^-$ we have

$$a^{-2\rho} \geq \sup_{\substack{v \in V, \\ \|v\|=1}} \langle \mathrm{Ad}(a)v, v \rangle.$$

We deduce an upper bound

$$\mathrm{vol}_Z(K\Omega_A aK \cdot z_0) \leq Ca^{-2\rho}. \tag{3.7}$$

We need to improve that bound from ρ to ρ_Q on the right hand side of (3.7). For that let W_L be the Weyl group of the reductive pair $(\mathfrak{l}, \mathfrak{a})$. Note that $\rho_Q =$

$\frac{1}{|W_L|}\sum_{w\in W_L} w\cdot\rho$. Further, the local structure theorem implies that $L_{\mathrm{n}}\subset H$ and hence W_L can be realized as a subgroup of $W_{H\cap K}:=N_{H\cap K}(\mathfrak{a})/Z_{H\cap K}(\mathfrak{a})$. We choose Ω_A to be invariant under $N_{H\cap K}(\mathfrak{a})$ and observe that $a\in A_Z$ is fixed under $W_{H\cap K}$. Thus using the $N_{H\cap K}(\mathfrak{a})$-symmetry in the a-variable we refine (3.7) to

$$\mathrm{vol}_Z(K\Omega_A aK\cdot z_0)\leq Ca^{-2\rho_Q}\,.$$

The desired bound then follows. □

Corollary 3.3. *Let $Z=G/H$ be a wavefront real spherical space of reductive type. Let $Z\to Z^\star$ be a basic factorization such that $Z^\star$ is not compact. Then the geometric balls B_R factorize well to $Z^\star$.*

Proof. As $Z\to Z^\star$ is basic we may assume (ignoring connected components) that $H^\star=G_IH$ for some I. Note that $\mathcal{F}=H^\star/H\simeq G_I/G_I\cap H$ is real spherical.
Let Q be the Z-adapted parabolic subgroup attached to P. Let $P_I=P\cap G_I$ and $G_I\supset Q_I\supset P_I$ be the $\mathcal{F}$-adapted parabolic above P_I and note that $Q_I=Q\cap G_I$. With Lemma 3.2 we then get

$$|B_R^{\mathcal{F}}|\asymp \sup_{\substack{X\in\mathfrak{a}_I\\ \|X\|\leq R}} e^{2\rho_{Q_I}(X)}\,,$$

which we are going to compare with (3.6).
Let $\mathfrak{u}_I$ be the Lie algebra of the unipotent radical of Q_I. Note that $\mathfrak{u}_I\subset\mathfrak{u}$ and that this inclusion is strict since $G/H^\star$ is not compact. The corollary now follows from (2.7). □

3.2 Property I

We briefly recall some results from [19].
Let $(\pi,\mathcal{H}_\pi)$ be a unitary irreducible representation of G. We denote by $\mathcal{H}_\pi^\infty$ the G-Fréchet module of smooth vectors and by $\mathcal{H}_\pi^{-\infty}$ its strong dual. One calls $\mathcal{H}_\pi^{-\infty}$ the G-module of distribution vectors; it is a DNF-space with continuous G-action.
Let $\eta\in(\mathcal{H}_\pi^{-\infty})^H$ be an H-fixed element and $H_\eta<G$ the stabilizer of η. Note that $H<H_\eta$ and set $Z_\eta:=G/H_\eta$. With regard to η and $v\in\mathcal{H}^\infty$ we form the generalized matrix-coefficient

$$m_{v,\eta}(gH):=\eta(\pi(g^{-1})v)\qquad(g\in G)$$

which is a smooth function on Z_η.
We recall the following facts from [19] Thm. 7.6 and Prop. 7.7:

Proposition 3.4. *Let Z be a wavefront real spherical space of reductive type. Then the following assertions hold:*

1. *Every generalized matrix coefficient $m_{v,\eta}$ as above is bounded.*

2. *Let $H < H^\star < G$ be a closed subgroup such that $Z^\star$ is unimodular. Then $Z^\star$ is a weakly basic factorization.*

3. *Let $(\pi, \mathcal{H})$ be a unitary irreducible representation of G and let $\eta \in (\mathcal{H}_\pi^{-\infty})^H$. Then:*

   (a) *$Z \to Z_\eta$ is a weakly basic factorization.*

   (b) *Z_η is unimodular and there exists $1 \leq p < \infty$ such that $m_{v,\eta} \in L^p(Z_\eta)$ for all $v \in \mathcal{H}_\pi^\infty$.*

The property of $Z = G/H$ that (3b) is valid for all π and η as above is denoted *Property (I)* in [19]. Note that (1) and (3b) together imply $m_{v,\eta} \in L^q(Z_\eta)$ for $q > p$. Assuming Property (I) we can then make the following notation.

Definition 3.5. *Given π as above, define $p_H(\pi)$ as the smallest index ≥ 1 such that all K-finite generalized matrix coefficients $m_{v,\eta}$ with $\eta \in (\mathcal{H}_\pi^{-\infty})^H$ belong to $L^p(Z_\eta)$ for any $p > p_H(\pi)$.*

Notice that $m_{v,\eta}$ belongs to $L^p(Z_\eta)$ for all K-finite vectors v once that this is the case for some non-trivial such vector v, see [19] Lemma 7.2. For example, this could be the trivial K-type, if it exists in π.
It follows from finite dimensionality of $(\mathcal{H}_\pi^{-\infty})^H$ (see [23]) that $p_H(\pi) < \infty$. We say that π is H-tempered if $p_H(\pi) = 2$.
The representation π is said to be H-distinguished if $(\mathcal{H}_\pi^{-\infty})^H \neq \{0\}$. Note that if π is not H-distinguished then $p_H(\pi) = 1$.

4 Lattice point counting: setup

Let G/H be a real algebraic homogeneous space. We further assume that we are given a lattice (a discrete subgroup with finite covolume) $\Gamma \subset G$, such that $\Gamma_H := \Gamma \cap H$ is a lattice in H. We normalize Haar measures on G and H such that:

- $\mathrm{vol}(G/\Gamma) = 1$.
- $\mathrm{vol}(H/\Gamma_H) = 1$.

Our concern is with the double fibration

$$\begin{tikzcd} & G/\Gamma_H \arrow[dl] \arrow[dr] & \\ Z := G/H & & Y := G/\Gamma \end{tikzcd}$$

Fibre-wise integration yields transfer maps from functions on Z to functions on Y and vice versa. In more precision,

$$L^\infty(Y) \to L^\infty(Z),\ \phi \mapsto \phi^H;\ \phi^H(gH) := \int_{H/\Gamma_H} \phi(gh\Gamma)\ d(h\Gamma_H) \tag{4.1}$$

and we record that this map is contractive, i.e

$$\|\phi^H\|_\infty \le \|\phi\|_\infty \qquad (\phi \in L^\infty(Y))\,. \tag{4.2}$$

Likewise we have

$$L^1(Z) \to L^1(Y),\ f \mapsto f^\Gamma;\ f^\Gamma(g\Gamma) := \sum_{\gamma\in\Gamma/\Gamma_H} f(g\gamma H)\,, \tag{4.3}$$

which is contractive, i.e

$$\|f^\Gamma\|_1 \le \|f\|_1 \qquad (f \in L^1(Z))\,. \tag{4.4}$$

Unfolding with respect to the double fibration yields, in view of our normalization of measures, the following adjointness relation:

$$\langle f^\Gamma, \phi\rangle_{L^2(Y)} = \langle f, \phi^H\rangle_{L^2(Z)} \tag{4.5}$$

for all $\phi \in L^\infty(Y)$ and $f \in L^1(Z)$. Let us note that (4.5) applied to $|f|$ and $\phi = \mathbf{1}_Y$ readily yields (4.4).
We write $\mathbf{1}_R \in L^1(Z)$ for the characteristic function of B_R and deduce from the definitions and (4.5):

- $\mathbf{1}_R^\Gamma(e\Gamma) = N_R(\Gamma, Z) := \#\{\gamma \in \Gamma/\Gamma_H \mid \gamma\cdot z_0 \in B_R\}$.
- $\|\mathbf{1}_R^\Gamma\|_{L^1(G/\Gamma)} = |B_R|$.

4.1 Weak asymptotics

In the above setup, G/H need not be of reductive type, but we shall assume this again from now on. For spaces with property (I) and Y compact we prove analytically in the following section that

$$N_R(\Gamma, Z) \sim |B_R| \quad (R \to \infty)\,. \tag{MT}$$

For that we will use the following result of [21]:

Theorem 4.1. *Let $Z = G/H$ be of reductive type. The smooth vectors for the regular representation of G on $L^p(Z)$ vanish at infinity, for all $1 \le p < \infty$.*

With notation from (4.3) we set

$$F_R^\Gamma := \frac{1}{|B_R|}\mathbf{1}_R^\Gamma.$$

We shall concentrate on verifying the following limit of weak type:

$$\langle F_R^\Gamma, \phi\rangle_{L^2(Y)} \to \int_Y \bar\phi\, d\mu_Y \quad (R\to\infty), \qquad (\forall \phi \in C_0(Y))\,. \tag{wMT}$$

Here C_0 indicates functions vanishing at infinity.

Lemma 4.2. (wMT) $\Rightarrow$ (MT).

Proof. As in [10] Lemma 2.3 this is deduced from Lemma 2.1 and Lemma 2.2. ∎

5 Main term counting

In this section we will establish main term counting under the mandate of property (I) and Y being compact. Let us call a family of balls $(B_R)_{R>0}$ *well factorizable* if it factorizes well to all proper factorizations of type $Z \to Z_\eta$.

5.1 Main theorem on counting

Theorem 5.1. *Let G be semi-simple and H a closed reductive subgroup. Suppose that Y is compact and Z admits* (I)*. If $(B_R)_{R>0}$ is well factorizable, then* (wMT) *and* (MT) *hold.*

Remark 5.2. In case $Z = G/H$ is real spherical and wavefront, then Z has (I) by Proposition 3.4. If we assume in addition that G has no compact factors and that all proper factorizations are basic, then the family of geometric balls is well factorizable by Corollary 3.3. In particular, Theorem A of the introduction then follows from the above.

The proof is based on the following proposition. For a function space $\mathcal{F}(Y)$ consisting of integrable functions on Y we denote by $\mathcal{F}(Y)_{\rm van}$ the subspace of functions with vanishing integral over Y.

Proposition 5.3. *Let $Z = G/H$ be of reductive type. Assume that there exists a dense subspace $\mathcal{A}(Y) \subset C_b(Y)^K_{\rm van}$ such that*

$$\phi^H \in C_0(Z) \quad \textit{for all } \phi \in \mathcal{A}(Y)\,. \tag{5.1}$$

Then (wMT) *holds true.*

Proof. We will establish (wMT) for $\phi \in C_b(Y)$. As

$$C_b(Y) = C_b(Y)_{\rm van} \oplus \mathbb{C}\mathbf{1}_Y,$$

and (wMT) is trivial for ϕ a constant, it suffices to establish

$$\langle F_R^\Gamma, \phi\rangle_{L^2(Y)} \to 0 \qquad (\phi \in C_b(Y)_{\rm van})\,. \tag{5.2}$$

We will show (5.2) is valid for $\phi \in \mathcal{A}(Y)$. By density, as F_R^Γ is K-invariant and belongs to $L^1(Y)$, this will finish the proof.
Let $\phi \in \mathcal{A}(Y)$ and let $\epsilon > 0$. By the unfolding identity (4.5) we have

$$\langle F_R^\Gamma, \phi\rangle_{L^2(Y)} = \frac{1}{|B_R|}\langle \mathbf{1}_R, \phi^H\rangle_{L^2(Z)}. \tag{5.3}$$

Using (5.1) we choose $K_\epsilon \subset Z$ compact such that $|\phi^H(z)| < \epsilon$ outside of K_ϵ. Then

$$\frac{1}{|B_R|}\langle \mathbf{1}_R, \phi^H\rangle_{L^2(Z)} = \int_{K_\epsilon} + \int_{Z-K_\epsilon} \frac{\mathbf{1}_R(z)}{|B_R|}\phi^H(z)\, d\mu_Z(z)\,.$$

By (4.2), the first term is bounded by $\frac{|K_\epsilon|\|\phi\|_\infty}{|B_R|}$, which is $\leq \epsilon$ for R sufficiently large. As the second term is bounded by ϵ for all R, we obtain (5.2). Hence (wMT) holds. ∎

Remark 5.4. It is possible to replace (5.1) by a weaker requirement: Suppose that an algebraic sum

$$\mathcal{A}(Y) = \sum_{j \in J} \mathcal{A}(Y)_j \tag{5.4}$$

is given together with a factorization $Z_j^\star = G/H_j^\star$ for each $j \in J$. Suppose that the balls B_R all factorize well to $Z_j^\star$, $j \in J$. Suppose further that ϕ^H factorizes to a function

$$\phi^{H_j^\star} \in C_0(Z_j^\star) \tag{5.5}$$

for all $\phi \in \mathcal{A}(Y)_j$ and all $j \in J$. Then the conclusion in Proposition 5.3 is still valid. In fact, using (2.5) the last part of the proof modifies to:

$$\begin{aligned} \frac{1}{|B_R|}\langle \mathbf{1}_R, \phi^H \rangle_{L^2(Z)} &= \frac{1}{|B_R|}\langle \mathbf{1}_R^{\mathcal{F}}, \phi^{H_j^\star} \rangle_{L^2(Z_j^\star)} = \\ &= \int_{K_\epsilon^\star} + \int_{Z_j^\star - K_\epsilon^\star} \frac{\mathbf{1}_R^{\mathcal{F}}(z)}{|B_R|} \phi^{H_j^\star}(z)\, d\mu_{Z_j^\star}(z) \end{aligned}$$

for $\phi \in \mathcal{A}(Y)_j$. As $\|1_R^{\mathcal{F}}\|_{L^1(Z_j^\star)} = |B_R|$, the second term is bounded by ϵ for all R. As the balls factorize well to $Z_j^\star$ we get the first term as small as we wish with (2.6).

5.2 The space $\mathcal{A}(Y)$

We now construct a specific subspace $\mathcal{A}(Y) \subset C_b(Y)_{\rm van}^K$ and verify condition (5.5).

Denote by $\widehat{G}_s \subset \widehat{G}$ the K-spherical unitary dual.

As Y is compact, the abstract Plancherel-theorem implies:

$$L^2(G/\Gamma)^K \simeq \bigoplus_{\pi \in \widehat{G}_s} (\mathcal{H}_\pi^{-\infty})^\Gamma.$$

If we denote the Fourier transform by $f \mapsto f^\wedge$ then the corresponding inversion formula is given by

$$f = \sum_\pi a_{v_\pi, f^\wedge(\pi)}. \tag{5.6}$$

Here $a_{v_\pi, f^\wedge(\pi)}$ denotes a matrix coefficient for Y with $v_\pi \in \mathcal{H}_\pi$ normalized K-fixed and $f^\wedge(\pi) \in (\mathcal{H}_\pi^{-\infty})^\Gamma$, and the sum in (5.6) is required to include multiplicities. The matrix coefficients for Y are defined as in (1.2), that is

$$a_{v,\nu}(y) = \nu(g^{-1} \cdot v) \qquad (y = gH \in Y)\,. \tag{5.7}$$

for $v \in \mathcal{H}_\pi$ and $\nu \in (\mathcal{H}_\pi^{-\infty})^\Gamma$.

Note that $L^2(Y) = L^2(Y)_{\rm van} \oplus \mathbb{C} \cdot \mathbf{1}_Y$. We define $\mathcal{A}(Y) \subset L^2(Y)_{\rm van}^K$ to be the dense subspace of functions with finite Fourier support, that is,

$$\mathcal{A}(Y) = \operatorname{span}\{a_{v,\nu} \mid \pi \in \widehat{G}_s \text{ non-trivial}, v \in \mathcal{H}_\pi^K, \nu \in (\mathcal{H}_\pi^{-\infty})^\Gamma\}.$$

Then $\mathcal{A}(Y) \subset L^2(Y)^{K,\infty}_{\mathrm{van}}$ is dense and since $C^\infty(Y)$ and $L^2(Y)^\infty$ are topologically isomorphic, it follows that $\mathcal{A}(Y)$ is dense in $C(Y)^K_{\mathrm{van}}$ as required.
The following lemma together with Remark 5.4 immediately implies Theorem 5.1.

Lemma 5.5. *Assume that Y is compact and Z has* (I)*, and define $\mathcal{A}(Y)$ as above. Then there exists a decomposition of $\mathcal{A}(Y)$ satisfying* (5.4)-(5.5)*.*

Proof. The map $\phi \mapsto \phi^H$ from (4.1) corresponds on the spectral side to a map $(\mathcal{H}_\pi^{-\infty})^\Gamma \to (\mathcal{H}_\pi^{-\infty})^H$, which can be constructed as follows.
As H/Γ_H is compact, we can define for each $\pi \in \widehat{G}_s$

$$\Lambda_\pi : (\mathcal{H}_\pi^{-\infty})^\Gamma \to (\mathcal{H}_\pi^{-\infty})^H, \quad \Lambda_\pi(\nu) = \int_{H/\Gamma_H} \nu \circ \pi(h^{-1})\ d(h\Gamma_H) \tag{5.8}$$

by $\mathcal{H}_\pi^{-\infty}$-valued integration: the defining integral is understood as integration over a compact fundamental domain $F \subset H$ with respect to the Haar measure on H; as the integrand is continuous and $\mathcal{H}_\pi^{-\infty}$ is a complete locally convex space, the integral converges in $\mathcal{H}_\pi^{-\infty}$. It follows from (5.8) that $(a_{v,\nu})^H = m_{v,\Lambda_\pi(\nu)}$ for all $v \in \mathcal{H}_\pi^\infty$ and $\nu \in (\mathcal{H}_\pi^{-\infty})^\Gamma$.
Let J denote the set of all factorizations $Z^\star \to Z$, including also $Z^\star = Z$ which we give the index $j_0 \in J$. For $j \in J$ we define $\mathcal{A}(Y)_j \subset \mathcal{A}(Y)$ accordingly to be spanned by the matrix coefficients $a_{v,\nu}$ for which $H_{\Lambda_\pi(\nu)} = H_j^\star$. Then (5.4) holds.
Let $\phi \in \mathcal{A}(Y)_{j_0}$, then it follows from (5.6) that

$$\phi^H = \sum_{\pi \neq \mathbf{1}} m_{v_\pi, \Lambda_\pi(\phi^\wedge(\pi))}\,. \tag{5.9}$$

Note that $H_\eta = H$ for each distribution vector $\eta = \Lambda_\pi(\phi^\wedge(\pi))$ in this sum, by the definition of $\mathcal{A}(Y)_{j_0}$. As Z has property (I) the summand $m_{v_\pi,\Lambda_\pi(\phi^\wedge(\pi))}$ is contained in $L^p(G/H)$ for $p > p_H(\pi)$, and by [19], Lemma 7.2, this containment is then valid for all K-finite generalized matrix coefficients $m_{v,\Lambda_\pi(\phi^\wedge(\pi))}$ of π. Thus $m_{v_\pi,\Lambda_\pi(\phi^\wedge(\pi))}$ generates a Harish-Chandra module inside $L^p(G/H)$. As $m_{v_\pi,\Lambda_\pi(\phi^\wedge(\pi))}$ is K-finite, we conclude that it is a smooth vector. Hence $\phi^H \in L^p(G/H)^\infty$, and in view of Theorem 4.1 we obtain (5.1).
The proof of (5.5) for $\phi \in \mathcal{A}(Y)_j$ for general $j \in J$ is obtained by the same reasoning, where one replaces H by $H_j^\star$ in (5.8) and (5.9). ◻

This concludes the proof of Theorem 5.1.

6 L^p-bounds for generalized matrix coefficients

¿From here on we assume that $Z = G/H$ is wavefront and real spherical. Recall that we assumed that G is semi-simple and that we wrote $\mathfrak{g} = \mathfrak{g}_1 \oplus \ldots \oplus \mathfrak{g}_m$ for the decomposition of $\mathfrak{g}$ into simple factors. It is no big loss of generality to

assume that $G = G_1 \times \ldots \times G_m$ splits accordingly. We will assume that from now on.
Further we request that the lattice $\Gamma < G$ is irreducible, that is, the projection of Γ to any normal subgroup $J \subsetneq G$ is dense in J.
Let π be an irreducible unitary representation of G. Then $\pi = \pi_1 \otimes \ldots \otimes \pi_m$ with π_j and irreducible unitary representation of G_j. We start with a simple observation.

Lemma 6.1. *Let $(\pi, \mathcal{H})$ be an irreducible unitary representation of G and $0 \neq \nu \in (\mathcal{H}^{-\infty})^\Gamma$. If one constituent π_j of π is trivial, then π is trivial.*

Proof. The element ν gives rise to a G-equivariant injection

$$\mathcal{H}^\infty \hookrightarrow C^\infty(Y), \quad v \mapsto (g\Gamma \mapsto \nu(\pi(g^{-1})v)) . \tag{6.1}$$

Say π_j is trivial and let $J := \prod_{\substack{i=1 \\ i\neq j}}^m G_i$. Let Γ_J be the projection of Γ to J. Then (6.1) gives rise to a J-equivariant injection $\mathcal{H}^\infty \hookrightarrow C^\infty(J/\Gamma_J)$. As Γ_J is dense in J, the assertion follows. □

We assume from now on that the cycle $H/\Gamma_H \subset Y$ is compact. This technical condition ensures that the vector valued average map (5.8) converges.

Lemma 6.2. *Let $(\pi, \mathcal{H})$ be a non-trivial irreducible unitary representation of G. Let $\nu \in (\mathcal{H}_\pi^{-\infty})^\Gamma$ such that $\eta := \Lambda_\pi(\nu) \in (\mathcal{H}_\pi^{-\infty})^H$ is non-zero. Then H_η/H is compact.*

Proof. Recall from Proposition 3.4 that $Z \to Z_\eta$ is weakly basic, and from Lemma 2.8 that then there exists $H \subset H_b \subset H_\eta$ such that H_η/H_b is compact and $Z \to Z_b$ is basic. Hence $\mathfrak{h}_b = \mathfrak{h}_I$ for some I. As π is irreducible it infinitesimally embeds into $C^\infty(Z_\eta)$ and hence also to $C^\infty(Z_b)$ on which G_i acts trivially for $i \in I$. It follows that π_i is trivial for $i \in I$. Hence Lemma 6.1 implies $I = \emptyset$ and thus $\mathfrak{h}_b = \mathfrak{h}$. □

In the sequel we use the Plancherel theorem (see [15])

$$L^2(G/\Gamma)^K \simeq \int_{\widehat{G}_s}^\oplus \mathcal{V}_{\pi,\Gamma}\ d\mu(\pi) ,$$

where $\mathcal{V}_{\pi,\Gamma} \subset (\mathcal{H}_\pi^{-\infty})^\Gamma$ is a finite dimensional subspace and of constant dimension on each connected component in the continuous spectrum (parametrization by Eisenstein series), and where the Plancherel measure μ has support

$$\widehat{G}_{\Gamma,s} := \operatorname{supp}(\mu) \subset \widehat{G}_s .$$

Given an irreducible lattice $\Gamma \subset G$ we define (cf. Definition 3.5)

$$p_H(\Gamma) := \sup\{p_H(\pi) : \pi \in \widehat{G}_{\Gamma,s}\} \tag{6.2}$$

and record the following.

Lemma 6.3. *Assume that $G = G_1 \times \ldots \times G_m$ with all $\mathfrak{g}_i$ simple and non-compact. Then $p_H(\Gamma) < \infty$.*

Proof. For a unitary representation $(\pi, \mathcal{H})$ and vectors $v, w \in \mathcal{H}$ we form the matrix coefficient $\pi_{v,w}(g) := \langle \pi(g)v, w\rangle$. We first claim that there exists a $p < \infty$ (in general depending on Γ) such that for all non-trivial $\pi \in \widehat{G}_{\Gamma,s}$ one has $\pi_{v,w} \in L^p(G)$ for all K-finite vectors v, w. In case G has property (T) this follows (independently of Γ) from [7]. The remaining cases contain at least one factor G_i of $\mathrm{SO}_e(n,1)$ or $\mathrm{SU}(n,1)$ (up to covering) and have no compact factors by assumption. They are treated in [6].
The claim can be interpreted geometrically via the leading exponent $\Lambda_V \in \mathfrak{a}^*$ which is attached to the Harish-Chandra module of $\mathcal{H}$ (see [19], Section 6). The lemma now follows from Prop. 4.2 and Thm. 6.3 in [19] (see the proof of Thm. 7.6 in [19] how these two facts combine to result in integrability). □

Let $1 \leq p < \infty$. Let us say that a subset $\Lambda \subset \widehat{G}_s$ is L^p*-bounded* provided that $m_{v,\eta} \in L^p(Z_\eta)$ for all $\pi \in \Lambda$ and $v \in \mathcal{H}_\pi^\infty$, $\eta \in (\mathcal{H}_\pi^{-\infty})^H$. By definition we thus have that $\widehat{G}_{\Gamma,s}$ is L^p-bounded for $p > p_H(\Gamma)$.
In this section we work under the following:

Hypothesis A: *For every $1 \leq p < \infty$ and every L^p-bounded subset $\Lambda \subset \widehat{G}_s$ there exists a compact subset $\Omega \subset G$ and constants $c, C > 0$ such that the following assertions hold for all $\pi \in \Lambda$, $\eta \in (\mathcal{H}_\pi^{-\infty})^H$ and $v \in \mathcal{H}_\pi^K$:*

$$\|m_{v,\eta}\|_{L^p(Z_\eta)} \leq C\|m_{v,\eta}\|_\infty \,, \tag{A1}$$

$$\|m_{v,\eta}\|_\infty \leq c\|m_{v,\eta}\|_{\infty,\Omega_\eta} \tag{A2}$$

where $\Omega_\eta = \Omega H_\eta/H_\eta$. Here $\|\cdot\|_{\infty,\omega}$ denotes the supremum norm taken on the subset ω.

In the sequel we are only interested in the following choice of subset $\Lambda \subset \widehat{G}_s$, namely

$$\Lambda := \{\pi \in \widehat{G}_{\Gamma,s} \mid \Lambda_\pi(\nu) \neq 0 \text{ for some } \nu \in \mathcal{V}_{\pi,\Gamma}\}\,. \tag{6.3}$$

An immediate consequence of Hypothesis A is:

Lemma 6.4. *Assume that $p > p_H(\Gamma)$. Then there is a $C > 0$ such that for all $\pi \in \widehat{G}_{\Gamma,s}$, $v \in \mathcal{H}_\pi^K$, $\nu \in (\mathcal{H}_\pi^{-\infty})^\Gamma$ and $\eta := \Lambda_\pi(\nu) \in (\mathcal{H}_\pi^{-\infty})^H$ one has*

$$\|\phi_\pi^H\|_{L^p(Z_\eta)} \leq C\|\phi_\pi\|_\infty$$

where $\phi_\pi(g\Gamma) := \nu(\pi(g^{-1})v)$.

Proof. Recall from (4.2), that integration is a bounded operator from $L^\infty(Y) \to L^\infty(Z)$. Hence the assertion follows from (A1). □

Recall the Cartan-Killing form κ on $\mathfrak{g} = \mathfrak{k} + \mathfrak{s}$ and choose a basis $X_1, \ldots, X_l$ of $\mathfrak{k}$ and $X_1', \ldots, X_s'$ of $\mathfrak{s}$ such that $\kappa(X_i, X_j) = -\delta_{ij}$ and $\kappa(X_i', X_j') = \delta_{ij}$. With that data we form the standard Casimir element

$$\mathcal{C} := -\sum_{j=1}^{l} X_j^2 + \sum_{j=1}^{s} (X_j')^2 \in \mathcal{U}(\mathfrak{g}) .$$

Set $\Delta_K := \sum_{j=1}^{l} X_j^2 \in \mathcal{U}(\mathfrak{k})$ and obtain the commonly used Laplace element

$$\Delta = \mathcal{C} + 2\Delta_K \in \mathcal{U}(\mathfrak{g}) \tag{6.4}$$

which acts on $Y = G/\Gamma$ from the left.
Let $d \in \mathbb{N}$. For $1 \leq p \leq \infty$, it follows from [2], Section 3, that Sobolev norms on $L^p(Y)^\infty \subset C^\infty(Y)$ can be defined by

$$||f||_{p,2d}^2 = \sum_{j=0}^{d} ||\Delta^j f||_p^2 .$$

Basic spectral theory allows one to define $\| \cdot \|_{p,d}$ more generally for any $d \geq 0$.
Let us define

$$s := \dim \mathfrak{s} = \dim G/K = \dim \Gamma \backslash G/K$$

and

$$r := \dim \mathfrak{a} = \operatorname{rank}_{\mathbb{R}}(G/K) ,$$

where $\mathfrak{a} \subset \mathfrak{s}$ is maximal abelian.
We denote by $C_b(Y)$ the space of continuous bounded functions on Y and by $C_b(Y)_{\mathrm{van}}$ the subspace with vanishing integral.

Proposition 6.5. *Assume that*

1. *Z is a wavefront real spherical space,*
2. *$G = G_1 \times \ldots \times G_m$ with all $\mathfrak{g}_i$ simple and non-compact.*
3. *$\Gamma < G$ is irreducible and Y_H is compact,*
4. *Hypothesis A is valid.*

Let $p > p_H(\Gamma)$. Then the map

$$\mathrm{Av_H} : C_b^\infty(Y)_{\mathrm{van}}^K \to L^p(Z)^K ; \mathrm{Av_H}(\phi) = \phi^H$$

is continuous. More precisely, for all

1. *$k > s + 1$ if Y is compact.*
2. *$k > \frac{r+1}{2} s + 1$ if Y is non-compact and Γ is arithmetic*

there exists a constant $C = C(p,k) > 0$ such that

$$\|\phi^H\|_{L^p(Z)} \leq C\|\phi\|_{\infty,k} \qquad (\phi \in C_b^\infty(Y)_{\rm van}^K)$$

Proof. For all $\pi \in \widehat{G}$ the operator $d\pi(\mathcal{C})$ acts as a scalar λ_π and we set

$$|\pi| := |\lambda_\pi| \geq 0\,.$$

Let $\phi \in C_b^\infty(Y)_{\rm van}^K$ and write $\phi = \phi_d + \phi_c$ for its decomposition in discrete and continuous Plancherel parts. We assume first that $\phi = \phi_d$.
In case Y is compact we have Weyl's law: There is a constant $c_Y > 0$ such that

$$\sum_{|\pi|\leq R} m(\pi) \sim c_Y R^{s/2} \qquad (R \to \infty)\,.$$

Here $m(\pi) = \dim \mathcal{V}_{\pi,\Gamma}$. We conclude that

$$\sum_\pi m(\pi)(1+|\pi|)^{-k} < \infty \tag{6.5}$$

for all $k > s/2+1$. In case Y is non-compact, we let $\widehat{G}_{\mu,d}$ be the the discrete support of the Plancherel measure. Then assuming Γ is arithmetic, the upper bound in [16] reads:

$$\sum_{\substack{\pi\in\widehat{G}_{\mu,d}\\ |\pi|\leq R}} m(\pi) \leq c_Y R^{rs/2} \qquad (R > 0)\,.$$

For $k > rs/2 + 1$ we obtain (6.5) as before.
Let $p > p_H(\Gamma)$. As ϕ is in the discrete spectrum we decompose it as $\phi = \sum_\pi \phi_\pi$ and obtain by Lemmas 6.2 and 6.4

$$\|\phi^H\|_p \leq \sum_\pi \|\phi_\pi^H\|_p \leq C\sum_\pi \|\phi_\pi\|_\infty\,.$$

The last sum we estimate as follows:

$$\begin{aligned}\sum_\pi \|\phi_\pi\|_\infty &= \sum_\pi (1+|\pi|)^{-k/2}(1+|\pi|)^{k/2}\|\phi_\pi\|_\infty\\ &\leq C\sum_\pi (1+|\pi|)^{-k/2}\|\phi_\pi\|_{\infty,k}\end{aligned}$$

with $C > 0$ a constant depending only on k (we allow universal positive constants to change from line to line). Applying the Cauchy-Schwartz inequality combined with (6.5) we obtain

$$\|\phi^H\|_p \leq C\Big(\sum_\pi \|\phi_\pi\|_{\infty,k}^2\Big)^{\frac{1}{2}}$$

with $C > 0$. With Hypothesis (A2) we get the further improvement:

$$\|\phi^H\|_p \leq C\Big(\sum_\pi \|\phi_\pi\|^2_{\Omega,\infty,k}\Big)^{\frac{1}{2}}$$

where the Sobolev norm is taken only over the compact set Ω.
To finish the proof we apply the Sobolev lemma on $K\backslash G$. Here Sobolev norms are defined by the central operator $\mathcal{C}$, whose action agrees with the left action of Δ. It follows that $\|f\|_{\infty,\Omega} \leq C\|f\|_{2,k_1,\Omega}$ with $k_1 > \frac{s}{2}$ for K-invariant functions f on G. This gives

$$\|\phi^H\|_p \leq C(\sum_\pi ||\phi_\pi||^2_{\Omega,2,k+k_1})^{\frac{1}{2}} = C||\phi||_{\Omega,2,k+k_1} \leq C||\phi||_{\infty,k+k_1}$$

which proves the proposition for the discrete spectrum.
If $\phi = \phi_c$ belongs to the continuous spectrum, where multiplicities are bounded (see [15]), the proof is simpler. Let μ_c be the restriction of the Plancherel measure to the continuous spectrum. As this is just Euclidean measure on r-dimensional space we have

$$\int_{\widehat{G}_s} (1 + |\pi|)^{-k} \, d\mu_c(\pi) < \infty \tag{6.6}$$

if $k > r/2$. We assume for simplicity in what follows that $m(\pi) = 1$ for almost all $\pi \in \operatorname{supp}\mu_c$. As $\sup_{\pi\in\operatorname{supp}\mu_c} m(\pi) < \infty$ the proof is easily adapted to the general case.
Let

$$\phi = \int_{\widehat{G}_s} \phi_\pi \, d\mu_c(\pi).$$

As $\|\phi^H\|_\infty \leq \|\phi\|_\infty$ we conclude with Lemma 6.4, (6.6) and Fubini's theorem that

$$\phi^H = \int_{\widehat{G}_s} \phi^H_\pi \, d\mu_c(\pi)$$

and, by the similar chain of inequalities as in the discrete case

$$\|\phi^H\|_p \leq C\|\phi\|_{\infty,k+k_1}$$

with $k > \frac{r}{2}$ and $k_1 > \frac{s}{2}$. This concludes the proof. □

7 Error term estimates

Recall $\mathbf{1}_R$, the characteristic function of B_R. The first error term for the lattice counting problem can be expressed by

$$\operatorname{err}(R,\Gamma) := \sup_{\substack{\phi\in C_b(Y)\\ \|\phi\|_\infty\leq 1}} |\left\langle \frac{\mathbf{1}^\Gamma_R}{|B_R|} - \mathbf{1}_Y, \phi \right\rangle| \qquad (R > 0),$$

and our goal is to give an upper bound for $\mathrm{err}(R,\Gamma)$ as a function of R. According to the decomposition $C_b(Y) = C_b(Y)_{\mathrm{van}} \oplus \mathbb{C}\mathbf{1}_Y$ we decompose functions as $\phi = \phi_o + \phi_1$ and obtain

$$\mathrm{err}(R,\Gamma) = \sup_{\substack{\phi\in C_b(Y)\\ \|\phi\|_\infty\le 1}} \frac{|\langle \mathbf{1}_R^\Gamma, \phi_o\rangle|}{|B_R|} = \sup_{\substack{\phi\in C_b(Y)\\ \|\phi\|_\infty\le 1}} \frac{|\langle \mathbf{1}_R, \phi_o^H\rangle|}{|B_R|}\,.$$

Further, from $\|\phi_o\|_\infty \le 2\|\phi\|_\infty$ we obtain that $\mathrm{err}(R,\Gamma) \le 2\,\mathrm{err}_1(R,\Gamma)$ with

$$\mathrm{err}_1(R,\Gamma) := \sup_{\substack{\phi\in C_b(Y)_{\mathrm{van}}\\ \|\phi\|_\infty\le 1}} \frac{|\langle \mathbf{1}_R^\Gamma, \phi\rangle|}{|B_R|} = \sup_{\substack{\phi\in C_b(Y)_{\mathrm{van}}\\ \|\phi\|_\infty\le 1}} \frac{|\langle \mathbf{1}_R, \phi^H\rangle|}{|B_R|}\,.$$

7.1 Smooth versus non-smooth counting

Like in the classical Gauss circle problem one obtains much better estimates for the remainder term if one uses a smooth cutoff. Let $\alpha \in C_c^\infty(G)$ be a non-negative test function with normalized integral. Set $\mathbf{1}_{R,\alpha} := \alpha * \mathbf{1}_R$ and define

$$\mathrm{err}_\alpha(R,\Gamma) := \sup_{\substack{\phi\in C_b(Y)_o^K\\ \|\phi\|_\infty\le 1}} \frac{|\langle \mathbf{1}_{R,\alpha}^\Gamma, \phi\rangle|}{|B_R|} = \sup_{\substack{\phi\in C_b(Y)_o^K\\ \|\phi\|_\infty\le 1}} \frac{|\langle \mathbf{1}_{R,\alpha}, \phi^H\rangle|}{|B_R|}\,.$$

Lemma 7.1. *Let $k > s+1$ if Y is compact and $k > \frac{r+1}{2}s + 1$ otherwise. Let $p > p_H(\Gamma)$ and q be such that $\frac{1}{p} + \frac{1}{q} = 1$. Then there exists $C > 0$ such that*

$$\mathrm{err}_\alpha(R,\Gamma) \le C\|\alpha\|_{1,k}|B_R|^{-\frac{1}{p}} \tag{7.1}$$

for all $R \ge 1$ and all $\alpha \in C_c^\infty(G)$.

Proof. First note that

$$\langle \mathbf{1}_{R,\alpha}, \phi^H\rangle = \langle \mathbf{1}_{R,\alpha}, (-\mathbf{1}+\Delta)^{k/2}(-\mathbf{1}+\Delta)^{-k/2}\phi^H\rangle\,.$$

With $\psi = (-\mathbf{1}+\Delta)^{-k/2}\phi$ we have $\|\psi\|_{\infty,k} \le C\|\phi\|_\infty$ for some $C > 0$. We thus obtain

$$\begin{aligned}\mathrm{err}_\alpha(R,\Gamma) &\le C \sup_{\substack{\psi\in C_b(Y)_o^K\\ \|\psi\|_{\infty,k}\le 1}} \frac{|\langle \mathbf{1}_{R,\alpha}, (-\mathbf{1}+\Delta)^{k/2}\psi^H\rangle|}{|B_R|}\\ &\le \frac{C}{|B_R|}\sup_{\substack{\psi\in C_b(Y)_o^K\\ \|\psi\|_{\infty,k}\le 1}} |\langle \mathbf{1}_{R,\alpha}, (-\mathbf{1}+\Delta)^{k/2}\psi^H\rangle|\end{aligned}$$

Moving $(-\mathbf{1}+\Delta)^{k/2}$ to the other side we get with Hölder's inequality and Proposition 6.5 that

$$\mathrm{err}_\alpha(R,\Gamma) \le \frac{C}{|B_R|}\|(-\mathbf{1}+\Delta)^{k/2}\alpha * \mathbf{1}_R\|_q\,.$$

Finally,

$$\|(-\mathbf{1}+\Delta)^{k/2}\alpha * \mathbf{1}_R\|_q \leq C\|\alpha\|_{1,k}\|\mathbf{1}_R\|_q$$

and with $\|\mathbf{1}_R\|_q = |B_R|^{\frac{1}{q}}$, the lemma follows. $\square$

Remark 7.2. In the literature results are sometimes stated not with respect to $\operatorname{err}(R,\Gamma)$ but the pointwise error term $\operatorname{err}_{pt}(R,\Gamma) = |\mathbf{1}_R^\Gamma(\mathbf{1}) - |B_R||$. Likewise we define $\operatorname{err}_{pt,\alpha}(R,\Gamma)$. Let B_Y be a compact neighborhood of $\mathbf{1}\Gamma \in Y$ and note that

$$\operatorname{err}_{pt,\alpha}(R,\Gamma) \leq |B_R| \sup_{\substack{\phi\in L^1(B_Y)\\ \|\phi\|_1\leq 1}} |\langle \frac{\mathbf{1}_{R,\alpha}^\Gamma}{|B_R|} - \mathbf{1}_Y, \phi\rangle| \qquad (R>0).$$

The Sobolev estimate $\|\phi\|_\infty \leq C\|\phi\|_{1,k}$, for K-invariant functions ϕ on B_Y and with $k = \dim Y/K$ the Sobolev shift, then relates these error terms:

$$\operatorname{err}_{pt,\alpha}(R,\Gamma) \leq |B_R| \sup_{\substack{\phi\in C_b^\infty(Y)\\ \|\phi\|_{\infty,-k}\leq 1}} |\langle \frac{\mathbf{1}_R^\Gamma}{|B_R|} - \mathbf{1}_Y, \phi\rangle|\,.$$

We then obtain

$$\operatorname{err}_{pt,\alpha}(R,\Gamma) \leq C|B_R|^{1-\frac{1}{p}} \qquad (R>0)$$

in view of (7.1).

We return to the error bound in Lemma 7.1 and would like to compare $\operatorname{err}_1(R,\Gamma)$ with $\operatorname{err}_\alpha(R,\Gamma)$. For that we note (by the triangle inequality) that

$$|\operatorname{err}_1(R,\Gamma) - \operatorname{err}_\alpha(R,\Gamma)| \leq \sup_{\substack{\phi\in C_b(Y)_o^K\\ \|\phi\|_\infty\leq 1}} \frac{|\langle \mathbf{1}_{R,\alpha}^\Gamma - \mathbf{1}_R^\Gamma, \phi\rangle|}{|B_R|}\,.$$

Suppose that $\operatorname{supp}\alpha \subset B_\epsilon^G$ for some $\epsilon > 0$. Then Lemma 2.2 implies that $\mathbf{1}_{R,\alpha}$ is supported in $B_{R+\epsilon}$, and hence

$$\begin{aligned} |\langle \mathbf{1}_{R,\alpha}^\Gamma - \mathbf{1}_R^\Gamma, \phi\rangle| &\leq \|\mathbf{1}_{R,\alpha}^\Gamma - \mathbf{1}_R^\Gamma\|_1 \\ &\leq \|\mathbf{1}_{R,\alpha} - \mathbf{1}_R\|_1 \\ &\leq |B_{R+\epsilon}|^{\frac{1}{2}}\|\mathbf{1}_{R,\alpha} - \mathbf{1}_R\|_2 \\ &\leq |B_{R+\epsilon}|^{\frac{1}{2}}|B_{R+\epsilon}\backslash B_R|^{\frac{1}{2}}\,. \end{aligned}$$

With Lemma 2.1 we get

$$|B_{R+\epsilon}\backslash B_R| \leq C\epsilon|B_R| \qquad (R\geq 1, \epsilon<1)\,.$$

Thus we obtain that

$$|\operatorname{err}_1(R,\Gamma) - \operatorname{err}_\alpha(R,\Gamma)| \leq C\epsilon^{\frac{1}{2}}\,.$$

Combining this with the estimate in Lemma 7.1 we arrive at the existence of $C > 0$ such that

$$\mathrm{err}_1(R, \Gamma) \leq C(\epsilon^{-k}|B_R|^{-\frac{1}{p}} + \epsilon^{\frac{1}{2}})$$

for all $R \geq 1$ and all $0 < \epsilon < 1$. The minimum of the function $\epsilon \mapsto \epsilon^{-k}c + \epsilon^{1/2}$ is attained at $\epsilon = (2kc)^{\frac{2}{2k+1}}$ and thus we get:

THEOREM 7.3. *Under the assumptions of Proposition 6.5 the first error term* $\mathrm{err}(R, \Gamma)$ *for the lattice counting problem on* $Z = G/H$ *can be estimated as follows: for all* $p > p_H(\Gamma)$ *and* $k > s + 1$ *for* Y *compact, resp.* $k > \frac{r+1}{2}s + 1$ *otherwise, there exists a constant* $C = C(p, k) > 0$ *such that*

$$\mathrm{err}(R, \Gamma) \leq C|B_R|^{-\frac{1}{(2k+1)p}}$$

for all $R \geq 1$.

REMARK 7.4. The point where we lose essential information is in the estimate (6.5) where we used Weyl's law. In the moment pointwise multiplicity bounds are available the estimate would improve. To compare the results with Selberg on the hyperbolic disc, let us assume that $p_H(\Gamma) = 2$. Then with $r = 1$ and $s = 2$ our bound is $\mathrm{err}(R, \Gamma) \leq C_\epsilon |B_R|^{-\frac{1}{14}+\epsilon}$ while Selberg showed $\mathrm{err}(R, \Gamma) \leq C_\epsilon |B_R|^{-\frac{1}{3}+\epsilon}$.

8 TRIPLE SPACES

In this section we verify our Hypothesis A for triple space $Z = G/H$ where $G = G' \times G' \times G'$, $H = \mathrm{diag}(G')$ and $G' = \mathrm{SO}_e(1, n)$ for some $n \geq 2$. Observe that $\mathrm{SO}_e(1, 2) \cong \mathrm{PSl}(2, \mathbb{R})$. We take $K' := \mathrm{SO}(n, \mathbb{R}) < G'$ as a maximal compact subgroup and set $K := K' \times K' \times K'$. Further we set $\mathfrak{s} := \mathfrak{s}' \times \mathfrak{s}' \times \mathfrak{s}'$. A maximal abelian subspace $\mathfrak{a} \subset \mathfrak{s}$ is then of the form

$$\mathfrak{a} = \mathfrak{a}_1' \times \mathfrak{a}_2' \times \mathfrak{a}_3'$$

with $\mathfrak{a}_i' \subset \mathfrak{s}'$ one dimensional subspaces. We recall the following result from [8].

PROPOSITION 8.1. *For the triple space the following assertion hold true:*

1. $G = KAH$ *if and only if* $\dim(\mathfrak{a}_1' + \mathfrak{a}_2' + \mathfrak{a}_3') = 2$.
2. *Suppose that all* $\mathfrak{a}_i'$ *are pairwise distinct. Then one has* PH *is open for all minimal parabolics* P *with Langlands-decomposition* $P = M_P A_P N_P$ *and* $A_P = A$.

We say that the choice of A is *generic* if all $\mathfrak{a}_i'$ are distinct and $\dim(\mathfrak{a}_1' + \mathfrak{a}_2' + \mathfrak{a}_3') = 2$.

The invariant measure dz on Z can then be estimated as

$$\int_Z f(z)\, dz \leq \int_K \int_A f(ka \cdot z_0) J(a) \;\; da\; dk \qquad (f \in C_c(Z), f \geq 0)$$

with

$$J(a) = \sup_{w \in W} a^{2w\rho} \tag{8.1}$$

by Lemma 3.2. Note that in this case the Weyl group W is just $\{\pm 1\}^3$.

8.1 Proof of the Hypothesis A

We first note that for all $\pi \in \widehat{G}_s$ the space of H-invariants

$$(\mathcal{H}_\pi^{-\infty})^H = \mathbb{C} I\,.$$

is one-dimensional, see [5], Thm. 3.1.
Write $\pi = \pi_1 \otimes \pi_2 \otimes \pi_3$ with each factor a K'-spherical unitary irreducible representation of G'. If we assume that $\pi \neq \mathbf{1}$ has non-trivial H-fixed distribution vectors, then at least two of the factors π_i are non-trivial.
Let v_i be normalized K'-fixed vectors of π_i and set $v = v_1 \otimes v_2 \otimes v_3$. Since Z is a multiplicity one space, the functional $I \in (\mathcal{H}_\pi^{-\infty})^H$ is unique up to scalars. Our concern is to obtain uniform L^p-bounds for the generalized matrix coefficients $f_\pi := m_{v,I}$:

$$f_\pi(g_1, g_2, g_3) := I(\pi_1(g_1)^{-1} v_1 \otimes \pi_2(g_2)^{-1} v_2 \otimes \pi_3(g_3)^{-1} v_3)\,,$$

when π belongs to the set Λ of (6.3).
We decompose $\Lambda = \Lambda_0 \cup \Lambda_1 \cup \{\mathbf{1}\}$ with $\Lambda_0 \subset \Lambda$ the set of $\pi \in \Lambda$ with all π_i non-trivial, and Λ_1 the set of π's with exactly one π_i to be trivial.
Consider first the case where $\pi \in \Lambda_1$, i.e. one π_i is trivial, say π_3. Then $\pi_2 = \pi_1^*$. We identify $Z \simeq G' \times G'$ via $(g,h) \mapsto (\mathbf{1}, g, h)H$ and obtain

$$f_\pi(g,h) = \langle \pi_1(g) v_1, v_1 \rangle\,,$$

a spherical function. Note that $Z_\eta \simeq G'$ and Hypothesis A follows from standard properties about K'-spherical functions on G'. To be more specific let $G' = N'A'K'$ be an Iwasawa-decomposition with middle-projection $\mathbf{a} : G' \to A'$, then

$$f_\pi(g,h) = \varphi_{\lambda_1}(g) := \int_{K'} \mathbf{a}(k'g)^{\lambda_1 - \rho'}\, dk'\,.$$

We use Harish-Chandra's estimates $|\varphi_\nu(a)| \leq a^\nu \varphi_0(a)$ and $\varphi_0(a) \leq C a^{-\rho}(1 + |\log a|)^d$ for $a \in A'$ in positive chamber. The condition of $\pi \in \Lambda_1$ implies that $\rho - \mathrm{Re}\,\lambda_1 > 0$ is bounded away from zero and Hypothesis A follows in this case.
Suppose now that $\pi \in \Lambda_0$, i.e. all π_i are non-trivial.
For a simplified exposition we assume that $n = 2$, i.e. $G' = \mathrm{PSl}(2, \mathbb{R})$, and comment at the end for the general case. Then $\pi_i = \pi_{\lambda_i}$ are principal series for some $\lambda_i \in i\mathbb{R}^+ \cup [0,1)$ with $\mathcal{H}_{\pi_i}^\infty = C^\infty(\mathbb{S}^1)$ in the compact realization. Set $\lambda = (\lambda_1, \lambda_2, \lambda_3)$ and set $\pi = \pi_\lambda$.

In order to analyze f_π we use $G = KAH$ and thus assume that $g = a = (a_1, a_2, a_3) \in A$. We work in the compact model of $\mathcal{H}_{\pi_i} = L^2(\mathbb{S}^1)$ and use the explicit model for I in [3]: for h_1, h_2, h_3 smooth functions on the circle one has

$$\begin{aligned} I(h_1 \otimes h_2 \otimes h_3) = & \frac{1}{(2\pi)^3} \int_0^{2\pi} \int_0^{2\pi} \int_0^{2\pi} h_1(\theta_1) h_2(\theta_2) h_3(\theta_3) \cdot \\ & \cdot \mathcal{K}(\theta_1, \theta_2, \theta_3) \; d\theta_1 d\theta_2 d\theta_3 \,, \end{aligned}$$

where

$$\mathcal{K}(\theta_1, \theta_2, \theta_3) = |\sin(\theta_2 - \theta_3)|^{(\alpha-1)/2} |\sin(\theta_1 - \theta_3)|^{(\beta-1)/2} |\sin(\theta_1 - \theta_2)|^{(\gamma-1)/2} .$$

In this formula one has $\alpha = \lambda_1 - \lambda_2 - \lambda_3$, $\beta = -\lambda_1 + \lambda_2 - \lambda_3$ and $\gamma = -\lambda_1 - \lambda_2 + \lambda_3$ where $\lambda_i \in i\mathbb{R} \cup (-1, 1)$ are the standard representation parameters of π_i. According to to [5], Cor. 2.1, the kernel $\mathcal{K}$ is absolutely integrable.
Set

$$A' := \left\{ a_t := \begin{pmatrix} t & 0 \\ 0 & \frac{1}{t} \end{pmatrix} \mid t > 0 \right\} < G'$$

Then $A'_i = k_{\phi_i} A' k_{\phi_i}^{-1}$ with $\phi_i \in [0, 2\pi]$ and

$$k_\phi = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix} .$$

Set $a_{t,i} = k_{\phi_i} a_t k_{\phi_i}^{-1}$.
Returning to our analysis of f_π we now take $h_i(t_i, \theta_i) = [\pi_1(a_{t_i,i}) v_i](\theta_i)$ and remark that

$$h_i(t_i, \theta_i) = \frac{1}{(t_i^2 + \sin^2(\theta_i - \phi_i)(\frac{1}{t_i^2} - t_i^2))^{\frac{1}{2}(1+\lambda_i)}} .$$

Let us set $|\pi| := \pi_{\operatorname{Re}\lambda_1} \otimes \pi_{\operatorname{Re}\lambda_2} \otimes \pi_{\operatorname{Re}\lambda_3}$. Our formulas then show

$$|f_\pi(a)| \leq f_{|\pi|}(a) \qquad (a \in A) . \tag{8.2}$$

Let $c_i := 1 - |\operatorname{Re}\lambda_i|$ for $i = 1, 2, 3$. The fundamental estimate in [22], Thm. 3.2, then yields a constant d, independent of π, and a constant $C = C(\pi) > 0$ such that for $a = (a_{t_1,1}, a_{t_2,2}, a_{t_3,3})$ one has

$$|f_\pi(a)| \leq C \frac{(1 + |\log t_1| + |\log t_2| + |\log t_3|)^d}{[\cosh\log t_1]^{c_1} \cdot [\cosh\log t_2]^{c_2} \cdot [\cosh\log t_3]^{c_3}} . \tag{8.3}$$

In view of (8.2) the constant $C(\pi)$ can be assumed to depend only on the distance of $\operatorname{Re}\lambda_i$ to the trivial representation. Looking at the integral representation of f_π with the kernel $\mathcal{K}$ we deduce a lower bound without the logarithmic factor, i.e. the bound is essentially sharp. Hence (8.1) together with the fact that all f_π for $\pi \in \Lambda_0$ are in $L^p(Z)$ for some $p < \infty$ implies that

$$\inf_{\pi\in\Lambda_0} c_i(\pi) > 0\,. \tag{8.4}$$

We now claim

$$\sup_{\pi\in\Lambda_0} \|f_\pi\|_p < \infty\,, \tag{8.5}$$

and

$$\sup_{\pi\in\Lambda_0} \|f_\pi\|_\infty < \infty\,. \tag{8.6}$$

For $0 < \epsilon < 1$ set $\Lambda_{\epsilon,\mathbb{R}} = [0, 1-\epsilon] \times [0, 1-\epsilon] \times [0, 1-\epsilon]$ and $\Lambda_\epsilon := i\mathfrak{a}^* + \Lambda_\epsilon$. It follows from (8.4) that there exists an $\epsilon > 0$ such that $\Lambda_0 \subset \Lambda_\epsilon$. We prove the stronger inequalities with Λ_0 replaced by Λ_ϵ. In view of (8.2) and (8.3) we may replace by Λ_ϵ by $\Lambda_{\epsilon,\mathbb{R}}$. Let $\mathcal{E}_\epsilon$ be the eight element set of extreme points of $\Lambda_{\epsilon,\mathbb{R}}$. For fixed $a = a_t$ and $\theta = (\theta_1, \theta_2, \theta_3)$ we let $F_\lambda(a, \theta) = \mathcal{K}(\theta)h_1(t_1, \theta_1)h_2(t_2, \theta_2)h_3(t_3, \theta_3)$ and note that the assignment $\Lambda_{\epsilon,\mathbb{R}} \to \mathbb{R}_+$, $\lambda \mapsto F_\lambda(a, \theta)$ is convex. Therefore we get for all $\lambda \in \Lambda_\epsilon$ that

$$f_\lambda(a) \le \sum_{\mu\in\mathcal{E}_\epsilon} f_\mu(a)\,.$$

In view of (8.3) the inequalities (8.5) and (8.6) then follow.
On the other hand for $g = \mathbf{1} = (\mathbf{1}, \mathbf{1}, \mathbf{1})$, the value $f_\pi(\mathbf{1})$ is obtained by applying I to the constant function $\mathbf{1} = \mathbf{1}\otimes\mathbf{1}\otimes\mathbf{1}$. This value has been computed explicitly by Bernstein and Reznikov in [3] as

$$\frac{\Gamma((\alpha+1)/4)\Gamma((\beta+1)/4)\Gamma((\gamma+1)/4)\Gamma((\delta+1)/4)}{\Gamma((1-\lambda_1)/2)\Gamma((1-\lambda_2)/2)\Gamma((1-\lambda_3)/2)}$$

where α, β, γ are as before and $\delta = -\lambda_1 - \lambda_2 - \lambda_3$. Stirling approximation,

$$|\Gamma(\sigma + it)| = \text{const.} e^{-\frac{\pi}{2}|t|}|t|^{\sigma-\frac{1}{2}}\left(1 + O(|t|^{-1})\right)$$

as $|t| \to \infty$ and σ is bounded, yields a lower bound for $f_\pi(\mathbf{1})$:

$$\inf_{\pi\in\Lambda_0} |f_\pi(\mathbf{1})| > 0\,. \tag{8.7}$$

As $\|f_\pi\|_\infty \ge |f_\pi(\mathbf{1})|$ the assertion (A1) of Hypothesis A is readily obtained from (8.5) and (8.7). Likewise (A2) with $\Omega = \{\mathbf{1}\}$ follows from (8.6) and (8.7). In general for $G' = \mathrm{SO}_e(1, n)$ one needs to compute the Bernstein-Reznikov integral. This was accomplished in [9].

Theorem 8.2. *Let $Z = G' \times G' \times G'/\operatorname{diag}(G')$ for $G' = \mathrm{SO}_e(1, n)$ and assume that H/Γ_H is compact. Then the first error term* $\mathrm{err}(R, \Gamma)$ *for the lattice counting problem on $Z = G/H$ can be estimated as follows: for all $p > p_H(\Gamma)$ there exists a $C = C(p) > 0$ such that*

$$\mathrm{err}(R, \Gamma) \le C|B_R|^{-\frac{1}{(6n+3)p}}$$

for all $R \ge 1$.

8.2 Cubic lattices

Here we let $G_0 = \mathrm{SO}_e(1,2)$ with the quadratic Q form defining G_0 having integer coefficients and anisotropic over $\mathbb{Q}$, for example

$$Q(x_0, x_1, x_2) = 2x_0^2 - 3x_1^2 - x_2^2 \,.$$

Then, according to Borel, $\Gamma_0 = G_0(\mathbb{Z})$ is a uniform lattice in G_0.
Next let k be a cubic Galois extension of $\mathbb{Q}$. Note that k is totally real. An example of k is the splitting field of the polynomial $f(x) = x^3 + x^2 - 2x - 1$. Let σ be a generator of the Galois group of $k|\mathbb{Q}$. Let $\mathcal{O}_k$ be the ring of algebraic integers of k. We define $\Gamma < G = G_0^3$ to be the image of $G_0(\mathcal{O}_k)$ under the embedding

$$G_0(\mathcal{O}_k) \ni \gamma \mapsto (\gamma, \gamma^\sigma, \gamma^{\sigma^2}) \in G \,.$$

Then $\Gamma < G$ is a uniform irreducible lattice with trace $H \cap \Gamma \simeq \Gamma_0$ a uniform lattice in $H \simeq G_0$.

9 Outlook

We discuss some topics of harmonic analysis on reductive homogeneous spaces which are currently open and would have immediate applications to lattice counting.

9.1 A conjecture which implies Hypothesis A

Hypothesis A falls in the context of a more general conjecture about the growth behavior of families of Harish-Chandra modules.
We let $Z = G/H$ be a real spherical space. Denote by $A_Z^- \subset A_Z$ the compression cone of Z (see Section 3) and recall that wavefront means that $A^- A_H / A_H = A_Z^-$ which, however, we do not assume for the moment.
We use V to denote Harish-Chandra modules for the pair $(\mathfrak{g}, K)$ and V^∞ for their unique moderate growth smooth Fréchet globalizations. These V^∞ are global objects in the sense that they are G-modules whereas V is defined in algebraic terms. We write $V^{-\infty}$ for the strong dual of V^∞. We say that V is H-distinguished provided that the space of H-invariants $(V^{-\infty})^H$ is non-trivial.
It is no big loss of generality to assume that A_Z^- is a sharp cone, as the edge of this cone is in the normalizer of H and in particular acts on the finite dimensional space of H-invariants.
As A_Z^- is pointed it is a fundamental domain for the little Weyl group and as such a simplicial cone (see [17], Section 9). If $\mathfrak{a}_Z^- = \log A_Z^-$, then we write $\omega_1, \ldots, \omega_r$ for a set of generators (spherical co-roots) of $\mathfrak{a}_Z^-$.
Set $\overline{Q} := \theta(Q)$ where θ is the Cartan involution determined by the choice of K. Note that $V/\overline{\mathfrak{q}}V$ is a finite dimensional $\overline{Q}$ module, in particular a finite

dimensional A_Z-module. Let $\Lambda_1, \ldots, \Lambda_N \in \mathfrak{a}_Z^*$ be the $\mathfrak{a}_{Z,\mathbb{C}}$-weight spectrum. Then we define the H-spherical exponent $\Lambda_V \in \mathfrak{a}_Z^*$ of V by

$$\Lambda_V(\omega_i) := \max_{1\leq j\leq N} \operatorname{Re} \Lambda_j(\omega_i)\,.$$

Further attached to V is a "logarithmic" exponent $d \in \mathbb{N}$. Having this data we recall the main bound from [22]

$$|m_{v,\eta}(a\cdot z_0)| \lessapprox a^{\Lambda_V}(1+\|\log a\|)^{d_V} \qquad (a \in A_Z^-)\,.$$

Conjecture 9.1. *Fix a K-type τ, a constant $C > 0$, and a compact subset $\Omega \subset G$. Then there exists a compact set $\Omega_A \subset A_Z^-$ such that for all Harish-Chandra modules V with $\|\Lambda_V\| \leq C$, all $v \in V[\tau]$ and all $\eta \in (V^{-\infty})^H$ one has*

$$\max_{\substack{a\in A_Z^- \\ g\in\Omega}} |m_{v,\eta}(ga\cdot z_0)| a^{-\Lambda_V}(1+\|\log a\|)^{-d_V} =$$
$$\max_{\substack{a\in \Omega_A \\ g\in\Omega}} |m_{v,\eta}(ga\cdot z_0)| a^{-\Lambda_V}(1+\|\log a\|)^{-d_V}\,.$$

It is easily seen that this conjecture implies Hypothesis A if all the generalized matrix coefficients $m_{v,\eta}$ are bounded, as for example it is the case when Z is wavefront (see Proposition 3.4(1)).

Remark 9.2. It might well be that a slightly stronger conjecture is true. For that we recall that a Harish-Chandra module V has a unique minimal globalization, the analytic model V^ω. The space V^ω is an increasing union of subspaces V_ϵ for $\epsilon \to 0$. The parameter ϵ parametrizes left G-invariant neigborhoods $\Xi_\epsilon \subset G_\mathbb{C}$ of $\mathbf{1}$ which decrease with $\epsilon \to 0$. Further V_ϵ consists of those vectors $v \in V^\omega$ for which the orbit map $G \to V^\omega, \ \ g \mapsto g \cdot v$ extends to a holomorphic map on Ξ_ϵ. For fixed $\epsilon, C > 0$ the strengthened conjecture would be that there exists a compact subset Ω_A such that for all Harish-Chandra modules V with $\|\Lambda_V\| \leq C$ and all $v \in V_\epsilon$ one has

$$\max_{a\in A_Z^-} |m_{v,\eta}(a\cdot z_0)| a^{-\Lambda_V}(1+\|\log a\|)^{-d_V} =$$
$$\max_{a\in \Omega_A} |m_{v,\eta}(a\cdot z_0)| a^{-\Lambda_V}(1+\|\log a\|)^{-d_V}\,.$$

Note that the compact set Ω is no longer needed, as $\Omega \cdot V_\epsilon \subset V_{\epsilon'}$.

9.2 Spectral geometry of Z_η

In the general context of a reductive real spherical space it may be possible to establish both main term counting and the error term bound, with the arguments presented here for wavefront spaces, provided the following two key questions allow affirmative answers.

In what follows $Z = G/H$ is a real reductive spherical space and V denotes an irreducible Harish-Chandra module and $\eta \in (V^{-\infty})^H$.

Question A: *Is H_η reductive?*

Question B: *If for $v \in V$ the generalized matrix coefficient $m_{v,\eta}$ is bounded, then there exists a $1 \leq p < \infty$ such that $m_{v,\eta} \in L^p(Z_\eta)$.*

In this context we note that issues related to the well-factorization of the intrinsic balls in affine spherical spaces can possibly be resolved with similar methods to those applied here, using volume estimates as described in Theorem 7.17 of [13].

Acknowledgement. We are grateful to an anonymous referee for valuable comments that improved the presentation in the present version of the paper.

References

[1] Y. Benoist and H. Oh, *Effective equidistribution of S-integral points on symmetric varieties*, Ann. Inst. Fourier (Grenoble) 62 (2012), no. 5, 1889–1942.

[2] J. Bernstein and B. Krötz, *Smooth Fréchet globalizations of Harish-Chandra modules*, Israel J. Math. 199 (2014), 35–111.

[3] J. Bernstein and A. Reznikov, *Estimates of automorphic functions*, Moscow Math. J. 4 (1) (2004), 19–37

[4] P. Bravi and G. Pezzini, *The spherical systems of the wonderful reductive subgroups*, J. Lie Theory 25 (2015), 105–123.

[5] J.-L. Clerc and B. Ørsted, *Conformally invariant trilinear forms on the sphere*. Ann. Inst. Fourier 61 (2011), 1807–1838.

[6] L. Clozel, *Démonstration de la conjecture τ*, Invent. Math. 151 (2003), no. 2, 297–328.

[7] M. Cowling, *Sur les coefficients des representations unitaires des groupes de Lie simples*, Lecture Notes in Mathematics 739 (1979), 132–178.

[8] T. Danielsen, B. Krötz and H. Schlichtkrull, *Decomposition theorems for triple spaces*, Geom. Dedicata 174 (2015), 145–154.

[9] A. Deitmar, *Invariant triple products*, Int. J. Math. Math. Sci. 2006, Art. ID 48274, 22 pp.

[10] W. Duke, Z. Rudnick and P. Sarnak, *Density of integer points on affine homogeneous varieties*, Duke Math. J. 71 (1993), no. 1, 143–179.

[11] A. Eskin and C. McMullen, *Mixing, counting, and equidistribution in Lie groups*, Duke Math. J. 71 (1993), no. 1, 181–209.

[12] A. Eskin, S. Mozes and N. Shah, *Unipotent flows and counting lattice points on homogeneous varieties*, Ann. of Math. (2) 143 (1996), 253–299.

[13] A. Gorodnik and A. Nevo, The ergodic theory of lattice points, Princeton University press 2008.

[14] —, *Counting lattice points*, J. Reine Angew. Math. 663 (2012), 127–176.

[15] Harish-Chandra, Automorphic Forms on Semisimple Lie Groups, Springer LNM 62.

[16] L. Ji, *The Weyl upper bound on the discrete spectrum of locally symmetric spaces*, J. Diff. Geom. 51 (1) (1999), 97–147.

[17] F. Knop and B. Krötz, *Reductive group actions*, arXiv: 1604.01005

[18] F. Knop, B. Krötz, E. Sayag and H. Schlichtkrull, *Simple compactifications and polar decomposition for real spherical spaces*, Selecta Math. (N.S.) 21 (2015), 1071–1097.

[19] —, *Volume growth, temperedness and integrability of matrix coefficients on a real spherical space*, arXiv: 1407.8006. J. Funct. Anal. 271 (1) (2016), 12–36.

[20] F. Knop, B. Krötz and H. Schlichtkrull, *The local structure theorem for real spherical varieties*, Compositio Math. 151 (2015), 2145–2159.

[21] B. Krötz, E. Sayag and H. Schlichtkrull, *Vanishing at infinity on homogeneous spaces of reductive type*, arXiv: 1211.2781. Compositio Math., to appear.

[22] —, *Decay of matrix coefficients on reductive homogeneous space of spherical type*, Math. Z. 278 (2014), 229–249.

[23] B. Krötz and H. Schlichtkrull, *Multiplicity bounds and the subrepresentation theorem for real spherical spaces*, Trans. Amer. Math. Soc. 368 (2016), 2749-2762.

[24] Y. Sakellaridis and A. Venkatesh, *Periods and harmonic analysis on spherical varieties*, arXiv:1203.0039

Universität Paderborn
Institut für Mathematik
Warburger Straße 100
33098 Paderborn
Deutschland
bkroetz@gmx.de

Department of Mathematics
Ben Gurion University
of the Negev
P.O.B. 653
Be'er Sheva 84105, Israel
eitan.sayag@gmail.com

University of Copenhagen
Department of Mathematics
Universitetsparken 5
DK-2100 Copenhagen Ø
Denmark
schlicht@math.ku.dk

Decomposable Cycles and Noether-Lefschetz Loci

Kieran G. O'Grady[1]

Received: June 10, 2015
Revised: December 14, 2015

Communicated by Gavril Farkas

Dedicato alla piccola Titti

Abstract. We prove that there exist smooth surfaces of degree d in $\mathbb{P}^3$ whose group of rational equivalence classes of decomposable 0-cycles has rank at least $\lfloor \frac{d-1}{3} \rfloor$.

2010 Mathematics Subject Classification: 14C15
Keywords and Phrases: Decomposable cycles, Noether-Lefschetz loci.

0. Introduction

Let X be a smooth complex surface: a rational equivalence class of 0-cycles on X is *decomposable* if it is the intersection of two divisor classes. Let $\mathrm{DCH}_0(X) \subset \mathrm{CH}_0(X)$ be the subgroup generated by decomposable 0-cycles. Beaville and Voisin [1] proved that if X is a $K3$ surface then $\mathrm{DCH}_0(X) \cong \mathbb{Z}$. What can be said of the group $\mathrm{DCH}_0(X)$ in general? An irregular surface X with non-zero map $\bigwedge^2 H^0(\Omega^1_X) \to H^0(\Omega^2_X)$ provides an example with group of decomposable 0-cycles that is not finitely generated, even after tensorization with $\mathbb{Q}$. Let us assume that X is a regular surface: then $\mathrm{DCH}_0(X)$ is finitely generated because $\mathrm{CH}^1(X)$ is finitely generated, and we may ask for its rank. Blowing up regular surfaces with non-zero geometric genus at $(r-1)$ very general points, one gets examples of regular surfaces with $\mathrm{DCH}_0(X)$ of rank at least r (see Example 1.3 b) of [2]). What about a less artificial class of surfaces, such as (smooth) surfaces in $\mathbb{P}^3$? If the rank of $\mathrm{DCH}_0(X)$ is to be larger than 1 then the rank of $\mathrm{CH}^1(X)$ must be larger than 1, but the latter condition is not sufficient, for example curves on X whose canonical line-bundle is a (fractional) power of the hyperplane bundle do not increase the rank of $\mathrm{DCH}_0(X)$, see Subsection 1.2. The papers [13, 4] provide examples of smooth surfaces in $\mathbb{P}^3$ with Picard group of large rank and generated by lines: it follows that the group spanned

[1]Partially supported by PRIN 2013, the Giorgio and Elena Petronio Fellowship Fund, the Giorgio and Elena Petronio Fellowship Fund II, the Fund for Mathematics of the IAS

by decomposable 0-cycles of such surfaces has rank 1. On the other hand Lie Fu proved that there exist degree-8 surfaces $X \subset \mathbb{P}^3$ such that $\mathrm{DCH}_0(X)$ has rank at least 2, see 1.4 of [6]. In the present paper we will prove the result below.

Theorem 0.1. *There exist smooth surfaces $X \subset \mathbb{P}^3$ of degree d such that the rank of $\mathrm{DCH}_0(X)$ is at least $\lfloor \frac{d-1}{3} \rfloor$.*

In particular the rank of the group of decomposable 0-cycles of a smooth surface in $\mathbb{P}^3$ can be arbitrarily large.

Let us explain the main ideas that go into the proof of Theorem 0.1. Let $C = C_1 \cup \ldots \cup C_n$ be the disjoint union of smooth irreducible curves $C_j \subset \mathbb{P}^3$. Suppose that $d \gg 0$, and that the curves C_j are not rationally canonical, i.e. there exists $e \in \mathbb{Z}$ such that $K_{C_j}^{\otimes m} \cong \mathscr{O}_{C_j}(e)$ only for $m = 0$; we prove that for a very general smooth $X \in |\mathscr{I}_C(d)|$, the classes $c_1(\mathscr{O}_X(1))^2, C_1 \cdot C_1, \ldots, C_n \cdot C_n$ in $\mathrm{CH}_0(X)$ are linearly independent. We argue as follows. Assume that they are *not* linearly independent for X very general; then there exists a non-zero $(a, r_1, \ldots, r_n) \in \mathbb{Z}^{n+1}$ such that

$$ac_1(\mathscr{O}_X(1))^2 + r_1 c_1(\mathscr{O}_X(C_1))^2 + \ldots + r_n c_1(\mathscr{O}_X(C_n))^2 = 0 \tag{0.1}$$

for all smooth $X \in |\mathscr{I}_C(d)|$. Now let $\pi \colon W \to \mathbb{P}^3$ be the blow up of C, let E be the exceptional divisor of π, and E_j be the component of E mapping to C_j. Let $\Lambda(d) := |\pi^* \mathscr{O}_{\mathbb{P}^3}(d)(-E)|$, and let $\mathscr{S} \subset W \times \Lambda(d)$ be the universal surface parametrized by $\Lambda(d)$. We let $p_W \colon \mathscr{S} \to W$ and $p_{\Lambda(d)} \colon \mathscr{S} \to \Lambda(d)$ be the projection maps. There is a natural identification $\Lambda(d) = |\mathscr{I}_C(d)|$, and the generic $S \in \Lambda(d)$ is isomorphic to the corresponding $X \in |\mathscr{I}_C(d)|$. Since (0.1) holds for all smooth X, an application of the spreading principle shows that the class

$$p_W^*(a\pi^* c_1(\mathscr{O}_{\mathbb{P}}^3(1))^2 + r_1 c_1(\mathscr{O}_W(E_1))^2 + \ldots + r_n c_1(\mathscr{O}_W(E_n))^2) \in \mathrm{CH}^2(\mathscr{S}) \tag{0.2}$$

is *vertical*, i.e. is represented by a linear combination of codimension-2 subvarieties $\Gamma_i \subset \mathscr{S}$ such that

$$\dim p_{\Lambda(d)}(\Gamma_i) < \dim \Gamma_i. \tag{0.3}$$

We prove that if the class in (0.2) is vertical, then $0 = a = r_1 = \ldots = r_n$. The key result that one needs is a Noether-Lefschetz Theorem for surfaces belonging to an integral codimension-1 closed subset $A \in \Lambda(d)$. More precisely one needs to prove that the following hold:

(1) If the generic $S \in A$ is isomorphic to $\pi(S) \subset \mathbb{P}^3$, i.e. S contains no fiber of $\pi \colon W \to \mathbb{P}^3$ over C, then $\mathrm{CH}^1(S)$ is generated (over $\mathbb{Q}$) by $\pi^* c_1(\mathscr{O}_{\mathbb{P}^3}(1))|_S, c_1(\mathscr{O}_S(E_1)), \ldots, c_1(\mathscr{O}_S(E_n))$.
(2) If the generic $S \in A$ contains a fiber R of $\pi \colon W \to \mathbb{P}^3$ over C, necessarily unique by genericity of S, then $\mathrm{CH}^1(S)$ is generated (over $\mathbb{Q}$) by the classes listed in Item (1), together with $c_1(\mathscr{O}_S(R))$.

The reason why such a Noether-Lefschetz Theorem is needed is the following. Let $\Gamma_i \subset \mathscr{S}$ be a codimension-2 subvariety such that (0.3) holds, and assume that the generic fiber of $\Gamma_i \to p_{\Lambda(d)}(\Gamma_i)$ has dimension 1; then $A := p_{\Lambda(d)}(\Gamma_i)$ is an integral closed codimension-1 subset of $\Lambda(d)$, and the restriction of Γ_i to the surface S_t parametrized by $t \in A$ is a divisor on S_t. Thus we are lead to prove the above Noether-Lefschetz result. There is a substantial literature on Noether-Lefschetz, but we have not found a result taylor made for our needs. A criterion of K. Joshi [9] is very efficient in disposing of "most" choices of a codimension-1 closed subset $A \in \Lambda(d)$. We deal with the remaining cases by appealing to the Griffiths-Harris approach to Noether-Lefschetz [8], as further developed by Lopez [12] and Brevik-Nollet [5].

The paper is organized as follows. In SECTION 1 we consider a smooth 3-fold V with trivial Chow groups, an ample divisor H on V and surfaces in the linear system $|\mathscr{I}_C(H)|$, where $C = C_1 \cup \ldots \cup C_n$ is the disjoint union of a fixed collection of smooth irreducible curves $C_i \subset V$. We prove that if the curves C_i are not rationally canonical, and a suitable Noether-Lefschetz Theorem holds, then the classes of $C_1^2, \ldots, C_n^2$ on a very general $X \in |\mathscr{I}_C(H)|$ are linearly independent, and they span a subgroup intersecting trivially the image of $\mathrm{CH}^2(V) \to \mathrm{CH}^2(X)$. In SECTION 2 we prove the required Noether-Lefschetz Theorem for $V = \mathbb{P}^3_{\mathbb{C}}$. In SECTION 3 we prove THEOREM 0.1 by combining the main results of SECTION 1 and SECTION 2.

CONVENTIONS AND NOTATION: We work over $\mathbb{C}$. Points are closed points.
Let X be a variety: "If x is a generic point of X, then..." is shorthand for "There exists an open dense $U \subset X$ such that if $x \in U$ then...". Similarly the expression "If x is a very general point of X, then..." is shorthand for "There exists a countable collection of closed nowhere dense $Y_i \in X$ such that if $x \in (X \setminus \bigcup_i Y_i)$ then...".
From now on we will denote by $\mathrm{CH}(X)$ the group of rational equivalence classes of cycles with *rational* coefficients. Thus if Z_1, Z_2 are cycles on X then $Z_1 \equiv Z_2$ means that for some non-zero integer ℓ the cycles ℓZ_1, ℓZ_2 are integral and rationally equivalent. If Z is a cycle on X we will often use the same symbol (i.e. Z) for the rational equivalence class represented by Z.

ACKNOWLEDGEMENTS: It is a pleasure to thank Angelo Lopez for useful exchanges on Noether-Lefschetz results.

1. THE FAMILY OF SURFACES CONTAINING GIVEN CURVES

1.1. THREEFOLDS WITH TRIVIAL CHOW GROUPS. Throughout the paper V is an integral smooth projective threefold.

HYPOTHESIS 1.1. *The cycle class map* $cl\colon \mathrm{CH}(V) \longrightarrow H(V;\mathbb{Q})$ *is an isomorphism.*

The archetypal such V is $\mathbb{P}^3$. A larger class of examples is given by 3-folds with an algebraic cellular decomposition (see Ex. 1.9.1 of [7]), and conjecturally the above assumption is equivalent to vanishing of $H^{p,q}(V)$ for $p \neq q$. An integral smooth projective threefold has *trivial* Chow group if HYPOTHESIS 1.1 holds.

Claim 1.2. *Let V be as above, in particular it has trivial Chow group. The natural map*

$$\mathrm{S}^2\,\mathrm{CH}^1(V) \longrightarrow \mathrm{CH}^2(V) \tag{1.1}$$

is surjective.

Proof. The natural map $\mathrm{S}^2\,H^2(V;\mathbb{Q}) \to H^4(V;\mathbb{Q})$ is surjective by Hard Lefschetz. The claim follows because of Hypothesis 1.1. $\square$

1.2. Standard relations. Let V be an integral smooth projective 3-fold with trivial Chow group. Let $X \subset V$ be a closed surface, and $i\colon X \hookrightarrow V$ be the inclusion map. Let $\mathscr{R}^s(X) \subset \mathrm{CH}^s(X)$ be the image of the restriction map

$$\begin{array}{ccc} \mathrm{CH}^s(V) & \longrightarrow & \mathrm{CH}^s(X) \\ \xi & \mapsto & i^*\xi \end{array} \tag{1.2}$$

Notice that $\mathscr{R}^2(X) \subset \mathrm{DCH}_0(X)$ by Claim 1.2. Suppose that $C \subset X$ is an integral smooth curve. We will assume that $C \cdot C$ makes sense in $\mathrm{CH}_0(X)$, for example that will be the case if X is $\mathbb{Q}$-factorial. We will list elements of the kernel of the map

$$\begin{array}{ccc} \mathscr{R}^2(X) \oplus \mathscr{R}^1(X) \oplus \mathscr{R}^0(X) & \longrightarrow & \mathrm{DCH}_0(X) \\ (\alpha,\beta,\gamma) & \mapsto & \alpha + C\cdot\beta + \gamma\cdot C\cdot C \end{array} \tag{1.3}$$

Let $j\colon C \hookrightarrow V$ be the inclusion map. By Cor. 8.1.1 of [7] the following relation holds in $\mathrm{CH}_0(X)$:

$$i^*(j_*[C]) = C \cdot c_1(\mathscr{N}_{X/V}) = C \cdot i^*\mathscr{O}_V(X). \tag{1.4}$$

Thus

$$\alpha_C - C \cdot i^*\mathscr{O}_V(X) = 0, \tag{1.5}$$

where $\alpha_C := i^*(j_*C) \in \mathscr{R}^2(X)$. Equation (1.5) is the *first standard relation.* Now suppose that there exists $\xi \in \mathrm{CH}^1(V)$ such that

$$c_1(K_C) = \xi|_C. \tag{1.6}$$

(Recall that Chow groups are with $\mathbb{Q}$-coefficients, thus (1.6) means that there exists an integer $n > 0$ such that $K_C^{\otimes n}$ is the pull-back of a line-bundle on V.) By adjunction for $X \subset V$ and for $C \subset X$,

$$C \cdot C + C \cdot (i^*K_V + i^*\mathscr{O}_X(X)) \equiv C \cdot i^*\xi. \tag{1.7}$$

Thus there exists $\beta_C \in \mathscr{R}^1(X)$ such that

$$\beta_C \cdot C - C \cdot C = 0. \tag{1.8}$$

The above is the *second standard relation* (it holds assuming (1.6)).

Example 1.3. Let $V = \mathbb{P}^3$, let $X \subset \mathbb{P}^3$ be a smooth surface of degree d, and let $C \subset X$ be a smooth curve. The subgroup of $\mathrm{CH}_0(X)$ spanned by intersections of linear combinations of $H := c_1(\mathscr{O}_X(1))$ and C has rank at most 2. In fact the first standard relation reads $dC \cdot H = (\deg C)H \cdot H$. Suppose that $c_1(K_C) = mC \cdot H$, where $m \in \mathbb{Q}$. With this hypothesis, the second standard relation reads $C \cdot C = (m + 4 - d)C \cdot H$, and hence $C \cdot C, C \cdot H, H \cdot H$ span a rank-1 subgroup. In particular a curve of genus 0 or 1 does not add anything to the rank of $\mathrm{DCH}_0(X)$.

1.3. SURFACES CONTAINING DISJOINT CURVES. Let V be a smooth projective 3-fold with trivial Chow group and $C_1,\ldots,C_n \subset V$ be pairwise disjoint integral smooth projective curves. Let $C := C_1 \cup \ldots \cup C_n$ and let $\pi\colon W \to V$ be the blow-up of C. Let E be the exceptional divisor of π, and let E_j, for $j \in \{1,\ldots,n\}$, be the irreducible component of E mapping to C_j. Let H be an ample divisor on V. For $j \in \{1,\ldots,n\}$ we let

(1.9)
$$\Sigma_j := \{S \in |\pi^*H - E| \mid \pi(S) \text{ is singular at some point of } C_j\}, \quad \Sigma := \cup_{j=1}^n \Sigma_j.$$

Let $S \in |\pi^*H - E|$, and let $X := \pi(S)$. Then $S \in \Sigma_j$ if and only if S contains one (at least) of the fibers of $E_j \to C_j$, or, equivalently, the map $S \to X$ given by restriction of π is *not* an isomorphism over C_j. We will always assume that $(\pi^*H - E)$ is very ample on W; with this hypothesis Σ_j is irreducible of codimension 1, or empty (compute the codimension of the loci of $S \in |\pi^*(H) - E|$ which contain one or two fixed fibers of $E_k \to C_k$). Suppose that H is sufficiently ample: then, in addition, if $S \in \Sigma_k$ is generic the surface $X = \pi(S)$ is smooth except for one ODP (ordinary double point) belonging to C_k, and the set of reducible $S \in |\pi^*H - E|$ is of large codimension in $|\pi^*H - E|$. We will assume that both of these facts hold (but we do not assume that H is "sufficiently ample", because we want to prove effective results).

HYPOTHESIS 1.4. *Let $C_1,\ldots,C_n \subset V$ and H be as above, in particular H is ample on V, and $(\pi^*H - E)$ is very ample on W. Suppose that*

(1) *for $j \in \{1,\ldots,n\}$, and $S \in \Sigma_j$ generic, the surface $\pi(S)$ is smooth except for one ODP (ordinary double point) belonging to C_j, and*
(2) *the set of reducible $S \in |\pi^*H - E|$ has codimension at least 3 in $|\pi^*H - E|$.*

Assume that HYPOTHESIS 1.4 holds, and let $S \in \Sigma_j$ be generic. Then there is a unique singular point of $\pi(S)$, call it x, and the line $\pi^{-1}(x)$ is contained in S.

HYPOTHESIS 1.5. *Let $C_1,\ldots,C_n \subset V$ and H be as above. Suppose that* HYPOTHESIS 1.4 *holds, and that in addition the following hold:*

(1) *If $A \subset |\pi^*H - E|$ is an integral closed codimension-1 subset, not equal to one of $\Sigma_1,\ldots,\Sigma_n$, and $S \in A$ is very general, the restriction map $\mathrm{CH}^1(W) \to \mathrm{CH}^1(S)$ is surjective.*
(2) *For $j \in \{1,\ldots,n\}$, $S \in \Sigma_j$ very general, and x the unique singular point of $\pi(S)$ (an ODP belonging to C_j, by* HYPOTHESIS 1.4*), $\mathrm{CH}^1(S)$ is generated by the image of the restriction map $\mathrm{CH}^1(W) \to \mathrm{CH}^1(S)$ together with the class of $\pi^{-1}(x)$.*

Remark 1.6. Let $V = \mathbb{P}^3$, and fix $C_1,\ldots,C_n \subset \mathbb{P}^3$. Let $d \gg 0$, and $H \in |\mathscr{O}_{\mathbb{P}^3}(d)|$. If $S \in \Sigma_j$ is generic, then $\pi^{-1}(x)$ does *not* belong to the image of the restriction map $\mathrm{CH}^1(W) \to \mathrm{CH}^1(S)$.

In the present section we will prove the following result.

PROPOSITION 1.7. *Let $C_1, \ldots, C_n \subset V$ and H be as above, and assume that* HYPOTHESIS 1.5 *holds. Suppose also that for $j \in \{1, \ldots, n\}$ there does* not *exist $\xi \in \mathrm{CH}^1(V)$ such that $c_1(K_{C_j}) = \xi|_{C_j}$. (Recall that Chow groups are with coefficients in $\mathbb{Q}$.) Then for very general smooth $X \in |\mathscr{I}_C(H)|$ the following hold:*

1. *The map $\mathrm{CH}^2(V) \to \mathrm{CH}_0(X)$ is injective.*
2. *Let $\{\zeta_1, \ldots, \zeta_m\}$ be a basis of $\mathrm{CH}^1(V)$ (as $\mathbb{Q}$-vector space). Suppose that for very general smooth $X \in |\mathscr{I}_C(H)|$*
$$0 = P(\zeta_1|X, \ldots, \zeta_m|X) + r_1 C_1^2 + \ldots + r_n C_n^2,$$
where $P \in \mathbb{Q}[x_1, \ldots, x_m]_2$ is a homogeneous quadratic polynomial. Then $0 = P(\zeta_1, \ldots, \zeta_m) = r_1 = \ldots = r_n$.

The proof of PROPOSITION 1.7 will be given in SUBSECTION 1.7. Throughout the present section we let V, C, W, E and H be as above.

1.4. THE UNIVERSAL SURFACE. Assume that HYPOTHESIS 1.4 holds. Let

$$\Lambda := |\pi^*(H) - E| \tag{1.10}$$
$$\mathscr{S} := \{(x, S) \in W \times \Lambda \mid x \in S\}. \tag{1.11}$$

Let $p_W : \mathscr{S} \to W$ and $p_\Lambda : \mathscr{S} \to \Lambda$ be the forgetful maps. Thus we have

$$\begin{tikzcd} & & \mathscr{S} \arrow[dl, "p_W"'] \arrow[dr, "p_\Lambda"] & \\ V & W \arrow[l, "\pi"'] & & \Lambda \end{tikzcd} \tag{1.12}$$

Let $N := \dim \Lambda$. Since $(\pi^*(H) - E)$ is very ample it is globally generated and hence the map p_W is a $\mathbb{P}^{N-1}$-fibration. It follows that $\mathscr{S}$ is smooth and

$$\dim \mathscr{S} = (N + 2). \tag{1.13}$$

DEFINITION 1.8. Let $\mathrm{Vert}^q(\mathscr{S}/\Lambda) \subset \mathrm{CH}^q(\mathscr{S})$ be the subspace spanned by rational equivalence classes of codimension-q integral closed subsets $Z \subset \mathscr{S}$ such that the dimension of $p_\Lambda(Z)$ is strictly smaller than the dimension of Z.

The result below is an instance of the spreading principle.

CLAIM 1.9. *Keep notation and assumptions as above, in particular* HYPOTHESIS 1.4 *holds. Let $Q \in \mathbb{Q}[x_1, \ldots, x_m, y_1, \ldots, y_n]_2$ be a homogeneous polynomial of degree 2 and let $\zeta_1, \ldots, \zeta_m \in \mathrm{CH}^1(V)$. Then*

$$Q(\zeta_1|_X, \ldots, \zeta_m|_X, c_1(\mathscr{O}_X(C_1)), \ldots, c_1(\mathscr{O}_X(C_n))) = 0 \tag{1.14}$$

for all smooth $X \in |\mathscr{I}_C(H)|$ if and only if

$$p_W^* Q(\pi^*\zeta_1, \ldots, \pi^*\zeta_m, E_1, \ldots, E_n) \in \mathrm{Vert}^2(\mathscr{S}/\Lambda). \tag{1.15}$$

Proof. Suppose that (1.14) holds for all smooth $X \in |\mathscr{I}_C(H)|$. Let $S \in \Lambda$ be generic, $X := \pi(S)$. Then X is smooth and the restriction of π to S defines an isomorphism $\varphi \colon S \xrightarrow{\sim} X$, thus by our assumption

$$p_W^* Q(\pi^*\zeta_1, \ldots, \pi^*\zeta_m, E_1, \ldots, E_n)|_S = 0.$$

Since S is generic in Λ it follows (see [3, 14]) that there exists an open dense subset $\mathscr{U} \subset \Lambda$ such that

$$p_W^* Q(\pi^*\zeta_1, \ldots, \pi^*\zeta_m, E_1, \ldots, E_n)|_{p_\Lambda^{-1}\mathscr{U}} = 0. \tag{1.16}$$

(We recall that Chow groups are with rational coefficients, if we consider integer coefficients then (1.16) holds only up to torsion.) Let $B := (\Lambda \setminus \mathscr{U})$. By the localization exact sequence

$$\mathrm{CH}_N(p_\Lambda^{-1}B) \longrightarrow \mathrm{CH}_N(\mathscr{S}) \longrightarrow \mathrm{CH}_N(p_\Lambda^{-1}\mathscr{U}) \longrightarrow 0$$

$p_W^* Q(\pi^*\zeta_1, \ldots, \pi^*\zeta_m, E_1, \ldots, E_n)$ is represented by an N-cycle supported on $p_\Lambda^{-1}B$, and hence (1.15) holds because $\dim B < N$. Next, suppose that (1.15) holds. Then, by definition, the left-hand side of (1.15) is represented by an N-cycle whose support is mapped by p_Λ to a proper closed subset $B \subset \Lambda$. Thus there exists an open dense $\mathscr{U} \subset \Lambda$ such that the restriction of $p_W^* Q(\pi^*\zeta_1, \ldots, \pi^*\zeta_m, E_1, \ldots, E_n)$ to $p_\Lambda^{-1}\mathscr{U}$ vanishes, e.g. $\mathscr{U} = \Lambda \setminus B$. By shrinking $\mathscr{U}$ we may assume that for $S \in \mathscr{U}$ the surface $X := \pi(S)$ is smooth. Let $S \in \mathscr{U}$: then $0 = p_W^* Q(\pi^*\zeta_1, \ldots, \pi^*\zeta_m, E_1, \ldots, E_n)|_S$, and since $X \cong S$ it follows that (1.14) holds for $X = \pi(S)$. On the other hand the locus of smooth $X \in |\mathscr{I}_C(H)|$ such that (1.14) holds is a countable union of closed subsets of Λ_{sm} (the open dense subset of Λ parametrizing smooth surfaces); since it contains an open dense subset of Λ_{sm} it is equal to Λ_{sm}. □

1.5. THE CHOW GROUPS OF $\mathscr{S}$ AND W. Assume that HYPOTHESIS 1.4 holds. Let $\xi \in \mathrm{CH}^1(\mathscr{S})$ be the pull-back of the hyperplane class on Λ via the map p_Λ of (1.12). Since p_W is the projectivization of a rank-N vector-bundle on W and ξ restricts to the hyperplane class on each fiber of p_W the Chow ring $\mathrm{CH}(\mathscr{S})$ is the $\mathbb{Q}$-algebra generated by $p_W^* \mathrm{CH}(W)$ and ξ, with ideal of relations generated by a single relation in codimension N. We have $N \geq 3$ because$(\pi^*H - E)$ is very ample by HYPOTHESIS 1.4; thus

$$\begin{array}{ccc} \mathbb{Q} \oplus \mathrm{CH}^1(W) \oplus \mathrm{CH}^2(W) & \xrightarrow{\sim} & \mathrm{CH}^2(\mathscr{S}) \\ (a_0, a_1, a_2) & \mapsto & a_0\xi^2 + p_W^*(a_1) \cdot \xi + p_W^*(a_2) \end{array} \tag{1.17}$$

is an isomorphism. The Chow groups $\mathrm{CH}_q(W)$ are computed by first describing $\mathrm{CH}_q(E_j)$ for $j \in \{1, \ldots, n\}$, and then considering the localization exact sequence

$$\bigoplus_j \mathrm{CH}_q(E_j) \longrightarrow \mathrm{CH}_q(W) \longrightarrow \mathrm{CH}_q(W \setminus (E_1 \cup \ldots \cup E_n)) \longrightarrow 0.$$

One gets an isomorphism

$$\begin{array}{ccc} \mathrm{CH}^1(V) \oplus \mathbb{Q}^n & \xrightarrow{\sim} & \mathrm{CH}^1(W) \\ (a, t_1, \ldots, t_n) & \mapsto & \pi^*a + \sum_{j=1}^n t_j E_j \end{array} \tag{1.18}$$

and an exact sequence

$$0 \longrightarrow \mathrm{CH}^2(W)_{\mathrm{hom}} \longrightarrow \mathrm{CH}^2(W) \xrightarrow{cl} H^4(W;\mathbb{Q}) \longrightarrow 0 \tag{1.19}$$

where $\mathrm{CH}^2(W)_{\mathrm{hom}}$ is described as follows. Let $\rho_j\colon E_j \to C_j$ be the restriction of the blow-up map π, and $\sigma_j\colon E_j \hookrightarrow W$ be the inclusion map; then we have an Abel-Jacobi isomorphism

$$\begin{array}{cccc} AJ\colon & \mathrm{CH}^2(W)_{\mathrm{hom}} & \xrightarrow{\sim} & \bigoplus_{j=1}^n \mathrm{CH}_0(C_j)_{\mathrm{hom}} \\ & \alpha & \mapsto & (\rho_{1,*}(\sigma_1^*\alpha), \ldots, \rho_{n,*}(\sigma_n^*\alpha)) \end{array} \tag{1.20}$$

Let AJ_j be the j-th component of the map AJ.

LEMMA 1.10. *Assume that* HYPOTHESIS 1.4 *holds. Let*

$$\omega := \pi^*\alpha + \sum_{j=1}^n E_j \cdot \pi^*\beta_j + \sum_{j=1}^n \gamma_j E_j \cdot E_j,$$

where $\alpha \in \mathrm{CH}^2(V)$, $\beta_j \in \mathrm{CH}^1(V)$, *and* $\gamma_j \in \mathbb{Q}$ *for* $j \in \{1, \ldots, n\}$. *Then the following hold:*

(1) *The cohomology class of* ω *vanishes if and only if*

$$\alpha = \sum_{j=1}^n \gamma_j C_j, \tag{1.21}$$

and for all $j \in \{1, \ldots, n\}$

$$\deg(\beta_j \cdot C_j) = -\gamma_j \deg(\mathscr{N}_{C_j/V}). \tag{1.22}$$

(2) *Suppose that* (1.21) *and* (1.22) *hold. Then for* $j \in \{1, \ldots, n\}$

$$AJ_j(\omega) = -\gamma_j c_1(\mathscr{N}_{C_j/V}) - c_1(\beta_j|_{C_j}). \tag{1.23}$$

Proof. Since the cohomology class map $cl\colon\ \mathrm{CH}^1(V) \to H^2(V;\mathbb{Q})$ is a surjection (by hypothesis), also the cohomology class map $cl\colon\ \mathrm{CH}^1(W) \to H^2(W;\mathbb{Q})$ is surjective. By Poincarè duality it follows that $cl(\omega) = 0$ if and only if $\deg(\omega \cdot \xi) = 0$ for all $\xi \in \mathrm{CH}^1(W)$. By (1.18) we must test $\xi = \pi^*\zeta$ with $\zeta \in \mathrm{CH}^1(V)$ and $\xi = E_i$ for $i \in \{1, \ldots, n\}$. We have

$$\deg(\omega \cdot \pi^*\zeta) = \deg\left(\left(\alpha - \sum_{j=1}^n \gamma_j C_j\right) \cdot \zeta\right). \tag{1.24}$$

Since the cycle map $\mathrm{CH}^2(V) \to H^4(V;\mathbb{Q})$ is an isomorphism, it follows that $\deg(\omega \cdot \pi^*\zeta) = 0$ for all $\zeta \in \mathrm{CH}^1(V)$ if and only if (1.21) holds. Next, we test $\xi = E_i$. In $\mathrm{CH}_0(C_i)$

$$\rho_{i,*} c_1(\mathscr{O}_{E_i}(E_i))^2 = -c_1(\mathscr{N}_{C_i/V}), \tag{1.25}$$

and hence

$$\deg(\omega \cdot E_i) = -\deg(\beta_i \cdot C_i) - \gamma_i \deg(\mathscr{N}_{C_i/V}). \tag{1.26}$$

This proves Item (1). Item (2) follows from Equation (1.25). □

Remark 1.11. By Lemma 1.10 the kernel of the map

$$\begin{array}{ccc} \mathrm{CH}^2(V)\oplus\bigoplus_{k=1}^{n}\mathrm{CH}^1(V)\oplus\bigoplus_{k=1}^{n}\mathbb{Q} & \longrightarrow & \mathrm{CH}^2(W) \\ (\alpha,\beta_1,\ldots\beta_n,\gamma_1,\ldots,\gamma_n) & \mapsto & \pi^*\alpha+\sum_{j=1}^n E_j\cdot\pi^*\beta_j+\sum_{j=1}^n\gamma_j E_j\cdot E_j \end{array} \tag{1.27}$$

is generated over $\mathbb{Q}$ by the classes $E_j\cdot\pi^*\beta$, where $\beta\in\mathrm{CH}^1(V)$ and $\beta|_{C_j}=0$, together with the classes

$$\pi^*[C_j]+E_j\cdot\pi^*\beta+E_j\cdot E_j, \tag{1.28}$$

where $\beta\in\mathrm{CH}^1(V)$, $\deg(\beta\cdot C_j)=-\deg(\mathscr{N}_{C_j/V})$, and

$$-c_1(\mathscr{N}_{C_j/V})-c_1(\beta|_{C_j})=0. \tag{1.29}$$

Next notice that (1.29) holds if and only if $c_1(K_{C_j})$ is equal to the restriction of a class in $\mathrm{CH}^1(V)$ i.e. (1.6) holds. Assume that this is the case, and that $X\in|\mathscr{I}_C(H)|$ is a surface smooth at all points of C_j. Let $S\in|\pi^*H-E|$ be the strict transform of S. Then S is isomorphic to X over C_j, and restricting to S the equation $\pi^*[C_j]+E_j\cdot\pi^*\beta+E_j\cdot E_j=0$ we get the second standard relation (1.8).

1.6. A vertical cycle on $\mathscr{S}$. According to Claim 1.9, for every codimension-2 relation that holds between $\mathscr{O}_X(C_1),\ldots,\mathscr{O}_X(C_n)$ and restrictions to X of divisors on V, where X is an arbitrary smooth member of $\in|\mathscr{I}_C(H)|$, there is a polynomial in classes of $\pi^*\mathrm{CH}^1(V)$ and the classes of the exceptional divisors of π which is "responsible" for the relation, i.e. when we pull-it back to $\mathscr{S}$ it is a vertical class. We have shown that $\pi^*[C_j]+E_j\cdot\pi^*\beta+E_j\cdot E_j$ is the class responsible for the second standard relation (1.8), see Remark 1.11, and in fact this class vanishes. In the present subsection we will write out a cycle responsible for the first standard relation (1.5), this time the pull-back to $\mathrm{CH}^2(\mathscr{S})$ is a non-zero vertical class. We record for later use the following formulae:

$$\sigma_{j,*}\rho_j^*c_1(\mathscr{N}_{C_j/V}) = \pi^*C_j+E_j\cdot E_j, \tag{1.30}$$

$$p_{W,*}(\xi^N) = (\pi^*H-E). \tag{1.31}$$

The first formula follows from the "Key formula" for π^*C_j, see Prop. 6.7 of [7]. The second formula is immediate (recall that $N=\dim\Lambda$). Let $j\in\{1,\ldots,n\}$. By Hypothesis 1.4 there exists an open dense $U\subset\Sigma_j$ such that, if $S\in U$, then $S\cdot E_j=\mathbf{L}_x+Z$, where $x\in C_j$ is the unique singular point of $\pi(S)$, $\mathbf{L}_x:=\pi^{-1}(x)$, and Z is the residual divisor (whose support does not contain $\mathbf{L}_x$). It follows that

$$E_j\cap p_\Lambda^{-1}(U)=\mathscr{V}_j+\mathscr{Z}_j, \tag{1.32}$$

where, for every $S\in U$, the restrictions to $E_j\cap S$ of $\mathscr{V}_j$, $\mathscr{Z}_j$ are equal to $\mathbf{L}_x$ and Z, respectively. We let

$$\Theta_j:=\overline{\mathscr{V}}_j. \tag{1.33}$$

Thus $p_\Lambda(\Theta_j) = \Sigma_j$, and the generic fiber of $\Theta_j \to \Sigma_j$ is a projective line. By Hypothesis 1.4 Θ_j is of pure codimension 2 in $\mathscr{S}$ (or empty), and hence

$$\Theta_j \in \mathrm{Vert}^2(\mathscr{S}/\Lambda). \tag{1.34}$$

The result below will be instrumental in writing out the class of Θ_j in $\mathrm{CH}^2(\mathscr{S})$ according to Decomposition (1.17).

Proposition 1.12. *Let $j \in \{1,\ldots,n\}$. Then*

$$p_{W,*}(\Theta_j \cdot \xi^{N-1}) = 2E_j \cdot \pi^* H - E_j \cdot E_j - \pi^* C_j. \tag{1.35}$$

Proof. Let $\alpha, \beta \in H^0(W, \pi^*(H) - E)$ be generic. Then $\mathrm{div}(\alpha|_{E_j})$ and $\mathrm{div}(\beta|_{E_j})$ are smooth divisors intersecting transversely at points $p_1, \ldots, p_s$. Let $q_i := \pi(p_i)$ for $i \in \{1,\ldots,s\}$. Let $R = \mathbb{P}(\langle \alpha, \beta \rangle) \subset \Lambda$; thus $p_\Lambda^{-1} R$ represents ξ^{N-1}. Given p_i, there exists $[\lambda_i, \mu_i] \in \mathbb{P}^1$ such that $\mathrm{div}(\lambda_i \alpha + \mu_i \beta)$ contains $\pi^{-1}(q_i)$, and hence $[\lambda_i \alpha + \mu_i \beta] \in R \cap \Sigma_j$. Conversely, every point of $R \cap \Sigma_j$ is of this type. The line R intersects transversely Σ_j because it is generic, and hence

$$p_{W,*}(\Theta_j \cdot \xi^{N-1}) = \sigma_{j,*}\rho_j^*(q_1 + \ldots + q_s). \tag{1.36}$$

Thus in order to compute $p_{W,*}(\Theta_j \cdot \xi^{N-1})$ we must determine the class of the 0-cycle $q_1 + \ldots + q_s$. Let $\phi\colon C_j \times R \to C_j$ and $\psi\colon C_j \times R \to R$ be the projections and $\mathscr{F}$ the rank-2 vector-bundle on $C_j \times R$ defined by

$$\mathscr{F} := \phi^*(\mathscr{N}^\vee_{C_j/V} \otimes \mathscr{O}_{C_j}(H)) \otimes \psi^* \mathscr{O}_R(1).$$

The composition of the natural maps

$$\langle \alpha, \beta \rangle \hookrightarrow H^0(W, \pi^* H - E) \longrightarrow H^0(E_j, \mathscr{O}_{E_j}(\pi^* H - E)) \longrightarrow H^0(C_j, \mathscr{N}^\vee_{C_j/V} \otimes \mathscr{O}_{C_j}(H)) \tag{1.37}$$

defines a section $\tau \in H^0(\mathscr{F})$ whose zero-locus consists of points $p_1', \ldots, p_s'$ such that $\pi(p_i') = q_i$. Now, the zero-locus of τ represents $c_2(\mathscr{F})$, and hence

$$p_{W,*}(\Theta_j \cdot \xi^{N-1}) = \sigma_{j,*}(\rho_j^*(\phi_* c_2(\mathscr{F})))$$

by (1.36). The formula

$$c_2(\mathscr{F}) = \phi^*(2c_1(\mathscr{O}_C(H)) - c_1(\mathscr{N}_{C/\mathbb{P}^3})) \cdot \psi^* c_1(\mathscr{O}_R(1)).$$

gives

$$p_{W,*}(\Theta_j \cdot \xi^{N-1}) = 2E_j \cdot \pi^* H - \sigma_{j,*}\left(\rho_j^* c_1(\mathscr{N}_{C_j/V})\right). \tag{1.38}$$

Then (1.35) follows from the above equality together with (1.30). □

Corollary 1.13. *Let $j \in \{1,\ldots,n\}$. Then*

$$\Theta_j = \xi \cdot p_W^* E_j + p_W^*(E_j \cdot \pi^* H - \pi^* C_j). \tag{1.39}$$

Proof. By (1.17) there exist $\beta_h \in \mathrm{CH}^h(W)$ for $h = 0, 1, 2$ such that

$$\Theta_j = \xi^2 \cdot p_W^* \beta_0 + \xi \cdot p_W^* \beta_1 + p_W^* \beta_2.$$

Restricting p_W to Θ_j we get a $\mathbb{P}^{N-2}$-fibration $\Theta_j \to E_j$: it follows that $\beta_0 = 0$ and $\beta_1 = E_j$. By (1.31)

$$p_{W,*}(\Theta_j \cdot \xi^{N-1}) = p_{W,*}(\xi^N \cdot p_W^* E_j + \xi^{N-1} \cdot p_W^* \beta_2) = (E_j \cdot \pi^* H - E_j \cdot E_j + \beta_2). \tag{1.40}$$

On the other hand $p_{W,*}(\Theta_j \cdot \xi^{N-1})$ is equal to the right-hand side of (1.35): equating that expression and the right-hand side of (1.40) we get $\beta_2 = (E_j \cdot \pi^* H - \pi^* C_j)$. $\square$

Corollary 1.14. *Let $j \in \{1, \ldots, n\}$. Then $p_W^*(E_j \cdot \pi^* H - \pi^* C_j) \in \mathrm{Vert}^2(\mathscr{S}/\Lambda)$.*

Proof. By Corollary 1.13 we have

$$p_W^*(E_j \cdot \pi^* H - \pi^* C_j) = \Theta_j - \xi \cdot p_W^* E_j.$$

Now $\Theta_j \in \mathrm{Vert}^2(\mathscr{S}/\Lambda)$ (see (1.34)) and $\xi \cdot p_W^* E_j \in \mathrm{Vert}^2(\mathscr{S}/\Lambda)$ because it is supported on the inverse image of a hyperplane via p_Λ; thus $p_W^*(E_j \cdot \pi^* H - \pi^* C_j) \in \mathrm{Vert}^2(\mathscr{S}/\Lambda)$. $\square$

By Claim 1.9 the relation $p_W^*(E_j \cdot \pi^* H - \pi^* C_j) \in \mathrm{Vert}^2(\mathscr{S}/\Lambda)$ gives a relation in $\mathrm{CH}(X)$ for an arbitrary smooth $X \in |\mathscr{I}_C(H)|$. In fact it gives the first standard relation (1.5).

1.7. Proof of the main result of the section.

Lemma 1.15. *Assume that* Hypothesis 1.5 *holds. Then the projection $\mathrm{CH}^2(\mathscr{S}) \to \mathrm{CH}^2(W)$ determined by (1.17) maps $\mathrm{Vert}^2(\mathscr{S}/\Lambda)$ to the subspace spanned by*

$$(E_1 \cdot \pi^* H - \pi^* C_1), \ldots, (E_j \cdot \pi^* H - \pi^* C_j), \ldots, (E_n \cdot \pi^* H - \pi^* C_n). \tag{1.41}$$

Proof. Let $Z \subset \mathscr{S}$ be an irreducible closed codimension-2 subset of $\mathscr{S}$ such that

$$\dim p_\Lambda(Z) < \dim Z = N. \tag{1.42}$$

Since the fibers of p_Λ are surfaces,

$$\dim p_\Lambda(Z) = \begin{cases} N-2, & \text{or} \\ N-1. \end{cases} \tag{1.43}$$

Suppose that $\dim p_\Lambda(Z) = N-2$. We claim that

$$Z = p_\Lambda^{-1}(p_\Lambda(Z)). \tag{1.44}$$

Since $Z \subset p_\Lambda^{-1}(p_\Lambda(Z))$, it will suffice to prove that $p_\Lambda^{-1}(p_\Lambda(Z))$ is irreducible of dimension N. First we notice that every irreducible component of $p_\Lambda^{-1}(p_\Lambda(Z))$ has dimension at least N. In fact, letting $\iota \colon p_\Lambda(Z) \hookrightarrow \Lambda$ be the inclusion and $\Delta_\Lambda \subset \Lambda \times \Lambda$ the diagonal, $p_\Lambda^{-1}(p_\Lambda(Z))$ is identified with $(\iota, p_\Lambda)^{-1}\Delta_\Lambda$, and the claim follows because Δ_Λ is a l.c.i. of codimension N. Since every fiber of p_Λ has dimension 2, it follows that every irreducible component of $p_\Lambda^{-1}(p_\Lambda(Z))$ dominates $p_\Lambda(Z)$. On the other hand, since $\mathrm{cod}(p_\Lambda(Z), \Lambda) = 2$, there exists an open dense $U \subset p_\Lambda(Z)$ such that $p_\Lambda^{-1}(t)$ is irreducible for all $t \in U$ by Hypothesis 1.4, and hence $p_\Lambda^{-1}(U)$ is irreducible of dimension N. It follows that there is a single irreducible component of $p_\Lambda^{-1}(p_\Lambda(Z))$ dominating $p_\Lambda(Z)$, and hence $p_\Lambda^{-1}(p_\Lambda(Z))$ is irreducible (of dimension N). We have proved (1.44). Since Λ is a projective space, $p_\Lambda([Z])$ is a multiple of $c_1(\mathscr{O}_\Lambda(1))^2$. It follows that the class of Z is a multiple of ξ^2 and hence the projection $\mathrm{CH}^2(\mathscr{S}) \to \mathrm{CH}^2(W)$ maps it to 0. Now assume that $\dim p_\Lambda(Z) = N-1$. Let $Y := p_\Lambda(Z)$. For $t \in \Lambda$, we let $S_t := p_\Lambda^{-1}(t)$. We distinguish between the two cases:

(1) $p_\Lambda(Z) \notin \{\Sigma_1, \ldots, \Sigma_n\}$.

(2) There exists $j \in \{1, \ldots, n\}$ such that $p_\Lambda(Z) = \Sigma_j$.

Suppose that (1) holds. Let $Y^{sm} \subset Y$ be the subset of smooth points. If $t \in Y^{sm}$, we may intersect the cycles Z and S_t in $p_\Lambda^{-1}(Y)$ (because S_t is a l.c.i.), and the resulting cycle class $Z \cdot S_t$ belongs to $\mathrm{CH}^1(S_t)$. By HYPOTHESIS 1.5 there exists $\Gamma \in \mathrm{CH}^1(W)$ such that $\Gamma|_{S_t} = Z \cdot S_t$ for $t \in Y^{sm}$. It follows that there exists an open dense $U \subset Y^{sm}$ such that

$$\Gamma|_{p_\Lambda^{-1}(U)} \equiv Z|_{p_\Lambda^{-1}(U)}.$$

(Recall that Chow groups are with $\mathbb{Q}$-coefficients.) By the localization sequence applied to $p_\Lambda^{-1}(U) \subset p_\Lambda^{-1}(Y)$, it follows that there exists a cycle $\Xi \in \mathrm{CH}_N(p_\Lambda^{-1}(Y \setminus U))$ such that

$$[Z] = \Xi + p_W^*(\Gamma) \cdot p_\Lambda^*([Y]).$$

Here, by abuse of notation, we mean cycle classes in $\mathrm{CH}_N(\mathscr{S})$: thus $[Z]$ and Ξ are actually the push-forwards of the corresponding classes in $\mathrm{CH}_N(p_\Lambda^{-1}(Y))$ and $\mathrm{CH}_N(p_\Lambda^{-1}(Y \setminus U))$ via the obvious closed embeddings. By (1.44) Ξ is represented by a linear combination of varieties $p_\Lambda^{-1}(B_i)$, where $B_1, \ldots, B_m$ are the irreducible components of $Y \setminus U$; it follows that $\Xi = a\xi^2$ for some $a \in \mathbb{Q}$. On the other hand $[Y] \in \mathrm{CH}^1(\Lambda) = \mathbb{Q}c_1(\mathscr{O}_\Lambda(1))$, and hence $p_W^*(\Gamma) \cdot p_\Lambda^*([Y]) = bp_W^*(\Gamma)\xi$ for some $b \in \mathbb{Q}$. It follows that the projection $\mathrm{CH}^2(\mathscr{S}) \to \mathrm{CH}^2(W)$ maps Z to 0. Lastly suppose that Item (2) holds. Arguing as above, one shows that there exist $\Gamma \in \mathrm{CH}^1(W)$, an open dense $U \subset Y$, a cycle $\Xi \in \mathrm{CH}_N(p_\Lambda^{-1}(Y \setminus U))$, and $a \in \mathbb{Q}$ such that

$$[Z] = \Xi + p_W^*(\Gamma) \cdot p_\Lambda^*([Y]) + a\Theta_j.$$

By COROLLARY 1.13 the projection $\mathrm{CH}^2(\mathscr{S}) \to \mathrm{CH}^2(W)$ maps $[Z]$ to $a(E_j \cdot \pi^*H - \pi^*C_j)$. This proves that $\mathrm{Vert}^2(\mathscr{S}/\Lambda)$ is mapped into the subspace spanned by the elements of (1.41). Since $[\Theta_j]$ is a vertical class and is mapped to $(E_j \cdot \pi^*H - \pi^*C_j)$, we have proved the lemma. $\square$

Proof of PROPOSITION 1.7. Let $P \in \mathbb{Q}[x_1, \ldots, x_m]$ be homogeneous of degree 2 and $r_1, \ldots, r_n \in \mathbb{Q}$. The set of smooth $X \in |\mathscr{I}_C(H)|$ such that

$$0 = P(\zeta_1|X, \ldots, \zeta_m|X) + r_1C_1^2 + \ldots + r_nC_n^2 \tag{1.45}$$

is a countable union of closed subsets of the open dense subset of $|\mathscr{I}_C(H)|$ parametrizing smooth surfaces. It follows that if the proposition is false then there exist P and $r_1, \ldots, r_n$, not all zero, such that (1.45) holds for all smooth $X \in |\mathscr{I}_C(H)|$. Now we argue by contradiction. By CLAIM 1.9

$$p_W^*(P(\pi^*\zeta_1, \ldots, \pi^*\zeta_m) + \sum_{j=1}^n r_jE_j^2) \in \mathrm{Vert}^2(\mathscr{S}/\Lambda). \tag{1.46}$$

By LEMMA 1.15 it follows that there exist rationals $s_1, \dots, s_n$ such that

$$P(\pi^*\zeta_1, \dots, \pi^*\zeta_m) + \sum_{j=1}^{n} r_j E_j^2 = \sum_{j=1}^{n} s_j (E_j \cdot \pi^* H - \pi^* C_j),$$

i.e.,

$$0 = \pi^*(P(\zeta_1, \dots, \zeta_m) + \sum_{j=1}^{n} s_j C_j) - \sum_{j=1}^{n} s_j E_j \cdot \pi^* H + \sum_{j=1}^{n} r_j E_j^2. \tag{1.47}$$

Let ω be the right hand side of (1.47); then the homology class of ω vanishes, and also the Abel-Jacobi image $AJ(\omega)$, notation as in (1.20). Item (2) of LEMMA 1.10, together with our hypothesis that there does *not* exist $\xi \in \mathrm{CH}^1(V)$ such that $c_1(K_{C_j}) = \xi|_{C_j}$, gives $r_j = 0$ for $j \in \{1, \dots, n\}$. By (1.21)

$$P(\zeta_1, \dots, \zeta_m) + \sum_{j=1}^{n} s_j C_j = 0, \tag{1.48}$$

and hence $\sum_{j=1}^{n} s_j E_j \cdot \pi^* H = 0$. Thus

$$0 = E_i \cdot \left(\sum_{j=1}^{n} s_j E_j \cdot \pi^* H \right) = -s_i \deg(C_i \cdot H). \tag{1.49}$$

for $i \in \{1, \dots, n\}$. By hypothesis H is ample, and hence $s_i = 0$ follows from (1.49). Thus $P(\zeta_1, \dots, \zeta_m) = 0$ by (1.48). □

2. NOETHER-LEFSCHETZ LOCI FOR LINEAR SYSTEMS OF SURFACES IN $\mathbb{P}^3$ WITH BASE-LOCUS

2.1. THE MAIN RESULT. In the present section we let $V = \mathbb{P}^3$. Thus $C_1, \dots, C_n \subset \mathbb{P}^3$, and $\pi \colon W \to \mathbb{P}^3$. We let $\Lambda(d) := |\pi^* \mathscr{O}_{\mathbb{P}^3}(d)(-E)|$. For $j \in \{1, \dots, n\}$ let $\Sigma_j(d) \subset \Lambda(d)$ be the subset Σ_j considered in SECTION 1; thus $\Sigma_j(d)$ parametrizes surfaces $S \in \Lambda(d)$ such that $\pi(S)$ is singular at some point of C_j. Let $\Sigma(d) := \Sigma_1(d) \cup \ldots \cup \Sigma_n(d)$. We denote the tangent sheaf of a smooth variety X by T_X. Below is the main result of the present section.

THEOREM 2.1. *Suppose that $d \geq 5$, and that the following hold:*

(1) *$\pi^* \mathscr{O}_{\mathbb{P}^3}(d-3)(-E)$ is very ample.*
(2) *$H^1(C, T_C(d-4)) = 0$.*
(3) *The sheaf $\mathscr{I}_C$ is $(d-2)$-regular.*
(4) *The curves $C_1, \dots, C_n$ are not planar.*

Then HYPOTHESIS 1.5 *holds for $H \in |\mathscr{O}_{\mathbb{P}^3}(d)|$.*

Recall that HYPOTHESIS 1.5 states that HYPOTHESIS 1.4 holds, and that Items (1) and (2) (our Noether-Lefschetz hypotheses) of HYPOTHESIS 1.5 hold. The proof that HYPOTHESIS 1.4 holds is elementary, and will be given in SUBSECTION 2.2. We will prove that Items (1) and (2) of HYPOTHESIS 1.5 hold by applying Joshi's main criterion (Prop. 3.1 of [9]), and the main idea in

Griffiths-Harris' proof of the classical Noether-Lefschetz Theorem [8], as further developed by Lopez [12] and Brevik-Nollet [5]. The proof will be outlined in SUBSECTION 2.3, details are in the remaining subsections.

Remark 2.2. Choose disjoint integral smooth curves $C_1, \ldots, C_n \subset \mathbb{P}^3$ such that for each $j \in \{1, \ldots, n\}$ there does *not* exist $r \in \mathbb{Q}$ such that $c_1(K_{C_j}) = rc_1(\mathscr{O}_{C_j}(1))$. Let $d \gg 0$. Then the hypotheses of THEOREM 2.1 are satisfied, and hence by PROPOSITION 1.7 the following holds: if $X \in |\mathscr{I}_C(d)|$ is very general, then the 0-cycle classes $c_1(\mathscr{O}_X(1))^2, c_1(\mathscr{O}_X(C_1))^2, \ldots, c_1(\mathscr{O}_X(C_n))^2$ are linearly independent. Thus the group of decomposable 0-cycles of X has rank at least $n+1$. The proof of THEOREM 0.1 is achieved by making the above argument effective, see SECTION 3.

2.2. DIMENSION COUNTS. We will prove that, if the hypotheses of THEOREM 2.1 are satisfied, then HYPOTHESIS 1.4 holds for $H \in |\mathscr{O}_{\mathbb{P}^3}(d)|$. First, H is ample on $\mathbb{P}^3$, and $\pi^*(H) - E$ is very ample on W because it is the tensor product of the line-bundle $\pi^*\mathscr{O}_{\mathbb{P}^3}(d-3)(-E)$, which is very ample by hypothesis, and the base-point free line-bundle $\pi^*\mathscr{O}_{\mathbb{P}^3}(3)$. Let $\Delta(r) \subset \Lambda(r)$ be the closed subset parametrizing singular surfaces.

PROPOSITION 2.3. *Suppose that $\pi^*\mathscr{O}_{\mathbb{P}^3}(r-1)(-E)$ is very ample. Then the following hold:*

(1) *Let $x \in C$. The linear system $|\mathscr{I}_x^2(r)| \cap |\mathscr{I}_C(r)|$ has base locus equal to C, and codimension 2 in $|\mathscr{I}_C(r)|$. If X is generic in $|\mathscr{I}_x^2(r)| \cap |\mathscr{I}_C(r)|$ then it has an ODP at x and no other singularity.*
(2) *Given $x \in W \setminus E$ there exists $S \in \Delta(r)$ which has an ODP at x and is smooth away from x.*
(3) *The closed subset $\Delta(r)$ is irreducible of codimension 1 in $\Lambda(r)$, and the generic $S \in \Delta(r)$ has a unique singular point, which is an ODP.*
(4) *Let $j \in \{1, \ldots, n\}$. If S is a generic element of $\Sigma_j(r)$, then $\pi(S)$ has a unique singular point x, which is an ODP (notice that S is smooth).*

Proof. Let $q \in \mathbb{P}^3 \setminus C$. Since $\pi^*\mathscr{O}_{\mathbb{P}^3}(r-1)(-E)$ is very ample there exists $X \in |\mathscr{I}_C(r-1)|$ such that $q \notin X$. Let $P \subset \mathbb{P}^3$ be a plane containing x but not q: then $X + P \in |\mathscr{I}_x^2(r)| \cap |\mathscr{I}_C(r)|$ does not pass through q, and this proves that $|\mathscr{I}_x^2(r)| \cap |\mathscr{I}_C(r)|$ has base locus equal to C. Since $\pi^*\mathscr{O}_{\mathbb{P}^3}(r-1)(-E)$ is very ample there exist $F, G \in H^0(\mathbb{P}^3, \mathscr{I}_C(r-1))$ and $q_1, \ldots, q_m \in (C \setminus \{x\})$ such that $V(F), V(G)$ are smooth and transverse at each point of $C \setminus \{q_1, \ldots, q_m\}$. Let $P \subset \mathbb{P}^3$ be a plane not passing through x: the pencil in $|\mathscr{I}_C(r)|$ spanned by $V(F) + P$ and $V(G) + P$ does not intersect $|\mathscr{I}_x^2(r)| \cap |\mathscr{I}_C(r)|$, and hence $|\mathscr{I}_x^2(r)| \cap |\mathscr{I}_C(r)|$ has codimension at least 2 in $|\mathscr{I}_C(r)|$. The codimension is equal to 2 because imposing on $X \in |\mathscr{I}_C(r)|$ that it be singular at $x \in C$ is equivalent to 2 linear equations being satisfied. In order to show that the singularities of a generic element of $|\mathscr{I}_x^2(r)| \cap |\mathscr{I}_C(r)|$ are as claimed we consider the embedding

$$\begin{array}{rccc} (2.1) & \mathbb{P}(H^0(\mathbb{P}^3, \mathscr{I}_x(1)) \oplus H^0(\mathbb{P}^3, \mathscr{I}_x(1))) & \longrightarrow & \Sigma_j(r) \\ & [A, B] & \mapsto & V(AF + BG) \end{array}$$

where F, G are as above. The image is a sublinear system of $|\mathscr{I}_x^2(r)| \cap |\mathscr{I}_C(r)|$ whose base locus is C, hence the generic $V(A \cdot F + B \cdot G)$ is smooth away from C by Bertini's Theorem. A local computation shows that the projectivized tangent cone of $V(AF + BG)$ at x is a smooth conic for generic A, B. Lastly let $q \in C \setminus \{x\}$. The set of $[A, B]$ such that $V(AF + BG)$ is singular at q has codimension 2 if $q \notin \{q_1, \ldots, q_m\}$, codimension 1 if $q \in \{q_1, \ldots, q_m\}$: it follows that for generic $[A, B]$ the surface $V(AF + BG)$ is smooth at all points of $C \setminus \{x\}$. This proves Item (1). The remaining items are proved similarly. □

Remark 2.4. Let $x \in C$. The proof of PROPOSITION 2.3 shows that the projectivized tangent cone at x of the generic $X \in |\mathscr{I}_x^2(r)| \cap |\mathscr{I}_C(r)|$ is the generic conic in $\mathbb{P}(T_x\mathbb{P}^3)$ containing the point $\mathbb{P}(T_xC)$.

PROPOSITION 2.5. *Suppose that $\pi^*\mathscr{O}_{\mathbb{P}^3}(r)(-E)$ is very ample and that $\pi^*\mathscr{O}_{\mathbb{P}^3}(r-3)(-E)$ is base point free. Then the locus of non-integral surfaces $S \in |\Lambda(r)|$ has codimension at least* 4.

Proof. Let $\mathrm{Dec}(r) \subset \Lambda(r)$ be the (closed) subset of non-integral surfaces, and $\mathrm{Dec}(r)_1, \ldots, \mathrm{Dec}(r)_m$ be its irreducible components. Let $j \in \{1, \ldots, m\}$; we will prove that the locus of non-integral surfaces $S \in \mathrm{Dec}(r)_j$ has codimension at least 4. Suppose first that the generic $S \in \mathrm{Dec}(r)_j$ contains one (at least) of the components of E, say E_k. Since $\pi^*\mathscr{O}_{\mathbb{P}^3}(r)(-E)$ is very ample, and E_k is a $\mathbb{P}^1$-bundle, the image of the restriction map

$$H^0(W, \pi^*\mathscr{O}_{\mathbb{P}^3}(r)(-E)) \to H^0(E_k, \pi^*\mathscr{O}_{\mathbb{P}^3}(r)(-E)|_{E_k})$$

has dimension at least 4, and hence the locus of $S \in |\pi^*\mathscr{O}_{\mathbb{P}^3}(r)(-E)|$ which contain E_k has codimension at least 4.

Next, suppose that the generic $S \in \mathrm{Dec}(r)_j$ does not contain any of the components of E. Let $\mathrm{Dec}(r)'_j \subset |\mathscr{I}_C(r)|$ be the image of $\mathrm{Dec}(r)_j$ under the natural isomorphism $\Lambda(r) \xrightarrow{\sim} |\mathscr{I}_C(r)|$. Let $X \in \mathrm{Dec}(r)'_j$ be generic; we claim that

$$\dim(\operatorname{sing} X \setminus C) \geq 1. \tag{2.2}$$

In fact $X = \pi(S)$, where $S \in \mathrm{Dec}(r)_j$ is generic, and since S is non-integral we may write $S = S_1 + S_2$ where S_1, S_2 are effective non-zero divisors on W (we will identify effective divisors and pure codimension-1 subschemes of W and $\mathbb{P}^3$). Thus $X = X_1 + X_2$, where $X_i := \pi(S_i)$. Since X_1, X_2 are effective non-zero divisors on $\mathbb{P}^3$ (non-zero because neither S_1 nor S_2 contains a component of E), their intersection has dimension at least 1. Now $X_1 \cap X_2 \subset \operatorname{sing} X$, hence in order to prove (2.2) it suffices to show that $X_1 \cap X_2$ is not contained in C. Suppose that $X_1 \cap X_2$ is contained in C; then, since it has dimension at least 1, there exists $k \in \{1, \ldots, n\}$ such that $X_1 \cap X_2$ contains C_k, and this implies that S contains E_k, contradicting our assumption. We have proved (2.2).

Next, let $p \neq q \in (\mathbb{P}^3 \setminus C)$, and let $\Omega_{p,q}(r) \subset |\mathscr{I}_C(r)|$ be the subset of divisors X which are singular at p, q, with degenerate quadratic terms. If $X \in \mathrm{Dec}(r)'_j$, then by (2.2) there exists a couple of distinct $p, q \in (X \setminus C)$ such that X is singular at p and q, with degenerate quadratic terms (in fact the set of such couples is infinite). Thus, if Item (2) holds, then

$$\mathrm{Dec}(r)'_j \subset \bigcup_{p \neq q \in (\mathbb{P}^3 \setminus C)} \Omega_{p,q}(r). \tag{2.3}$$

Hence it suffices to prove that the codimension of $\Omega_{p,q}$ in $|\mathscr{I}_C(r)|$ is 10 (as expected) for each $p \neq q \in (\mathbb{P}^3 \setminus C)$. Let $Y \in |\mathscr{I}_C(r-3)|$ be a surface not containing p nor q (it exists because $\pi^*\mathscr{O}_{\mathbb{P}^3}(r-3)(-E)$ is base point free), and consider the subset

$$P_Y := \{Y + Z \mid Z \in |\mathscr{O}_{\mathbb{P}^3}(3)|\}.$$

An explicit computation shows that the codimension of the set of $Z \in |\mathscr{O}_{\mathbb{P}^3}(3)|$ singular at p, q, with degenerate quadratic terms, has codimension 10: it follows that $\Omega_{p,q}(r) \cap P_Y$ has codimension 10, and hence $\Omega_{p,q}(r)$ has codimension 10 in $|\mathscr{I}_C(r)|$. □

PROPOSITION 2.3 and PROPOSITION 2.5 prove that, if the hypotheses of THEOREM 2.1 are satisfied, then HYPOTHESIS 1.4 holds for $H \in |\mathscr{O}_{\mathbb{P}^3}(d)|$.

2.3. OUTLINE OF THE PROOF THAT THE NOETHER-LEFSCHETZ HYPOTHESIS HOLDS. Let A be an integral closed codimension-1 subset of $\Lambda(d)$. Let $A^\vee \subset \Lambda(d)^\vee$ be the projective dual of A, i.e. the closure of the locus of projective tangent hyperplanes $\mathbf{T}_S A$ for S a point in the smooth locus A^{sm} of A. Since $\pi^*\mathscr{O}_{\mathbb{P}^3}(d)(-E)$ is very ample we have the natural embedding $W \hookrightarrow \Lambda(d)^\vee$, and hence it makes sense to distinguish between the following two cases:

(I) $A^\vee$ is not contained in W.
(II) $A^\vee$ is contained in W.

Thus (I) holds if and only if, for the generic $S \in A^{sm}$, the projective tangent hyperplane $\mathbf{T}_S A$ is a base point free linear subsystem of $\Lambda(d)$. On the other hand, examples of codimension-1 subsets of $\Lambda(d)$ for which (II) holds are given by $\Delta(d)$ and by $\Sigma_j(d)$ for $j \in \{1, \ldots, n\}$. In fact $\Delta(d)^\vee = W$ and $\Sigma_j(d)^\vee = E_j$. The last equality holds because $S \in \Lambda(d)$ belongs to $\Sigma_j(d)$ if and only if it is tangent to E_j, thus $\Sigma_j(d) = E_j^\vee$, and hence $\Sigma_j(d)^\vee = E_j$ by projective duality. Let $\mathrm{NL}(\Lambda(d) \setminus \Delta(d))$ be the *Noether-Lefschetz locus*, i.e. the set of those smooth surfaces $S \in \Lambda(d)$ such that the restriction map $\mathrm{Pic}(W)_{\mathbb{Q}} \to \mathrm{Pic}(S)_{\mathbb{Q}}$ is *not* surjective. As is well-known $\mathrm{NL}(\Lambda(d) \setminus \Delta(d))$ is a countable union of closed subsets of $\Lambda(d) \setminus \Delta(d)$. In SUBSECTION 2.5 we will apply Joshi's criterion (Proposition 3.1 of [9]) in order to prove the following result.

PROPOSITION 2.6. *Suppose that $d \geq 5$ and that the following hold:*

(1) *$\pi^*\mathscr{O}_{\mathbb{P}^3}(d)(-E)$ is ample.*
(2) *$H^1(C, T_C(d-4)) = 0$.*
(3) *The sheaf $\mathscr{I}_C$ (on $\mathbb{P}^3$) is $(d-2)$-regular.*

Let $A \subset \Lambda(d)$ be an integral closed subset of codimension 1, and suppose that there exists $S \in (A \setminus \Delta(d))$ such that A is smooth at S, and the projective tangent space $\mathbf{T}_S A$ is a base-point free linear subsystem of $\Lambda(d)$. Then $A \setminus \Delta(d)$ does not belong to the Noether-Lefschetz locus $NL(\Lambda(d) \setminus \Delta(d))$.

The above result deals with codimension-1 subsets $A \subset \Lambda(d)$ for which (I) above holds. Thus, in order to finish the proof of Theorem 2.1, it remains to deal with those A such that (II) above holds.

Definition 2.7. Given $p \in W$ and $F \subset T_pW$ a vector subspace, we let

$$\Lambda_{p,F}(d) := \{S \in |\mathscr{I}_p \otimes \pi^*\mathscr{O}_{\mathbb{P}^3}(d)(-E)| : F \subset T_pS\}. \tag{2.4}$$

Let $\Gamma(d) := |\mathscr{I}_C(d)|$. We have a tautological identification $\Lambda(d) \xrightarrow{\sim} \Gamma(d)$: we let $\Gamma_{p,F}(d) \subset \Gamma(d)$ be the image of $\Lambda_{p,F}(d)$, and for $j \in \{1,\dots,n\}$ we let $\Pi_j(d) \subset \Gamma(d)$ be the image of $\Sigma_j(d)$.

Notice that $\Lambda_{p,F}(d)$ and $\Gamma_{p,F}(d)$ are linear subsystems of $\Lambda(d)$ and $\Gamma(d)$ respectively. In Subsection 2.6 we will prove the result below by applying an idea of Griffiths-Harris [8] as further developed by Lopez [12] and Brevik-Nollet [5].

Proposition 2.8. *Suppose that the following hold:*

(1) *$d \geq 4$ and $\pi^*\mathscr{O}_{\mathbb{P}^3}(d-3)(-E)$ is very ample.*
(2) *None of the curves $C_1,\dots,C_n$ is planar.*

Let X be a very general element

(a) *of $\Gamma_{p,F}(d)$, where either $p \notin E$, or else $p \in E$ and*

$$T_p(\pi^{-1}(\pi(p))) \not\subset F \subsetneq T_pE, \tag{2.5}$$

(b) *or of $\Pi_j(d)$ for some $j \in \{1,\dots,n\}$.*

Then the Chow group $\mathrm{CH}^1(X)_{\mathbb{Q}}$ is generated by $c_1(\mathscr{O}_X(1))$ and the classes of $C_1,\dots,C_n$.

Granting Proposition 2.8, let us prove that the statement of Theorem 2.1 holds for A such that $A^\vee$ is contained in W. We will distinguish between the following two cases:

(IIa) $A \notin \{\Sigma_1(d),\dots,\Sigma_n(d)\}$.
(IIb) $A \in \{\Sigma_1(d),\dots,\Sigma_n(d)\}$.

Suppose that (IIa) holds. By projective duality A is the closure of

$$\bigcup_{p\in(A^\vee)^{sm}} \Lambda_{p,T_pA^\vee} \tag{2.6}$$

Let $p \in (A^\vee)^{sm}$ be generic. We claim that Item (a) of Proposition 2.8 hold for p and $F = T_pA^\vee$. In fact if $A^\vee \not\subset E$ then $p \notin E$ by genericity. If $A^\vee \subset E$ then $A^\vee$ is contained in E_j for a certain $j \in \{1,\dots,n\}$. We claim that $A^\vee$ is a proper subset of E_j, and it is not equal to a fiber of the restriction of π to E_j. In fact, if $A^\vee = E_j$, then $A = E_j^\vee = \Sigma_j(d)$, and that contradicts the assumption that(IIa) holds. Now suppose that $A^\vee = \pi^{-1}(q)$ for a certain $q \in C_j$. Let $S \in A$ be generic. Since A is the closure of (2.6), S is tangent to $\pi^{-1}(q)$, and hence contains $\pi^{-1}(q)$ because S has degree 1 on every fiber of $E_j \to C_j$. It follows that S is tangent to E_j, and hence $A \subset E_j^\vee = \Sigma_j(d)$, contradicting the hypothesis that (IIa) holds.

Thus Item (a) of Proposition 2.8 hold for $p \in (A^\vee)^{sm}$ generic and $F = T_p A^\vee$, and hence if $S \in \Lambda_{p,T_p A^\vee}(d)$ is very general, then $\mathrm{CH}^1(X)_{\mathbb{Q}}$ is generated by $c_1(\mathscr{O}_X(1))$ and the classes of $C_1, \ldots, C_n$.
On the other hand, since $A \not\subset \Sigma(d)$, S intersects transversely E, and hence the restriction of π to S is an isomorphism $S \xrightarrow{\sim} X$. It follows that $\mathrm{CH}^1(S)_{\mathbb{Q}}$ is equal to the image of $\mathrm{CH}^1(W)_{\mathbb{Q}} \to \mathrm{CH}^1(S)_{\mathbb{Q}}$. This proves that there exists $S \in A$ such that $\mathrm{CH}^1(S)_{\mathbb{Q}}$ is equal to the image of $\mathrm{CH}^1(W)_{\mathbb{Q}} \to \mathrm{CH}^1(S)_{\mathbb{Q}}$. Actually our argument proves that there exists such an S which is smooth if $A \neq \Delta(d)$, and that if $A = \Delta(d)$ there exists such an S whose singular set consists of a single ODP. If the former holds, then we are done because $\mathrm{NL}(A \setminus \Delta(d))$ is a countable union of closed subsets of $A \setminus \Delta(d)$, and we have shown that the complement is non-empty. If the latter holds, let $\Delta(d)^0 \subset \Delta(d)$ be the open dense subset parametrizing surfaces with an ODP and no other singular point, then the set of $S \in \Delta(d)^0$ such that $\mathrm{CH}^1(W) \to \mathrm{CH}^1(S)$ is *not* surjective is a countable union of closed subsets of $\Delta(d)^0$ (take a simultaneous resolution), and we are done because we have shown that the complement is non empty.
Lastly suppose that (IIb) holds, i.e. $A = \Sigma_j(d)$ for a certain $j \in \{1, \ldots, n\}$. By Proposition 2.3 there exists an open dense subset $\Sigma_j(d)^0 \subset \Sigma_j(d)$ with the following property. If $S \in \Sigma_j(d)^0$ and $X = \pi(S)$, then X has a unique singular point, call it x (obviously $x \in C_j$), which is an ordinary node, and the restriction of π to S is the blow-up of X with center x (in particular S is smooth). Now suppose that $S \in \Sigma_j(d)^0$ is very general. Then by Proposition 2.8 the Chow group $\mathrm{CH}^1(S)_{\mathbb{Q}}$ is generated by the image of $\mathrm{CH}^1(W)_{\mathbb{Q}} \to \mathrm{CH}^1(S)_{\mathbb{Q}}$ and the class of $\pi^{-1}(x)$. Now notice that the set of $S \in \Sigma_j(d)^0$ such that $\mathrm{CH}^1(S)$ is *not* generated by the image of $\mathrm{CH}^1(W)_{\mathbb{Q}}$ together with $\pi^{-1}(x)$ is a countable union of closed subsets of $\Sigma_j(d)^0$; since the complement is not empty, we are done. $\square$
Summing up: we have shown that in order to prove Theorem 2.1 it suffices to prove Proposition 2.6 and Proposition 2.8. The proofs are in the following subsections.

2.4. Infinitesimal Noether-Lefschetz results. We will recall a key result of K. Joshi. Let $U \subset H^0(W, \pi^*\mathscr{O}_{\mathbb{P}^3}(d)(-E))$ be a subspace and $\sigma \in U$ be non-zero. We let $S := V(\sigma)$, and we assume that S is smooth. Let $\mathfrak{m}_{\sigma,U} \subset \mathscr{O}_{\mathbb{P}(U),[\sigma]}$ be the maximal ideal and $\mathscr{T}_{\sigma,U} := \mathrm{Spec}(\mathscr{O}_{\mathbb{P}(U),[\sigma]}/\mathfrak{m}_\sigma^2)$ be the first-order infinitesimal neighborhood of $[\sigma]$ in $\mathbb{P}(U)$. We let $\mathscr{S}_{\sigma,U} \to \mathscr{T}_{\sigma,U}$ be the restriction of the family $\mathscr{S}_\Lambda \to \Lambda$ to $\mathscr{T}_{\sigma,U}$. The Infinitesimal Noether Lefschetz (INL) Theorem is valid at (U, σ) (see Section 2 of [9]) if the group of line-bundles on $\mathscr{S}_{\sigma,U}$ is equal to the image of the composition

$$\mathrm{Pic}(W) \longrightarrow \mathrm{Pic}(W \times_{\mathbb{C}} \mathscr{T}_{\sigma,U}) \longrightarrow \mathrm{Pic}(\mathscr{S}_{\sigma,U}). \tag{2.7}$$

Let $A \subset \Lambda(d)$ be an integral closed subset. Let $[\sigma]$ be a smooth point of A, and suppose that $S = V(\sigma)$ is smooth. Let $\mathbb{P}(U)$ be the projective tangent space

to A at $[\sigma]$. If the INL Theorem holds for (U,σ) then $A\setminus\Delta(d)$ does *not* belong to the Noether-Lefschetz locus $NL(\Lambda(d)\setminus\Delta(d))$.
Joshi [9] gave a cohomological condition which suffices for the validity of the INL Theorem. Suppose that $U\subset H^0(W,\pi^*\mathscr{O}_{\mathbb{P}^3}(d)(-E))$ generates $\pi^*\mathscr{O}_{\mathbb{P}^3}(d)(-E)$; we let $M(U)$ be the locally-free sheaf on W fitting into the exact sequence

$$0\longrightarrow M(U)\longrightarrow U\otimes\mathscr{O}_W\longrightarrow\pi^*\mathscr{O}_{\mathbb{P}^3}(d)(-E)\longrightarrow 0. \tag{2.8}$$

Proposition 2.9 (K. Joshi, Prop. 3.1 of [9]). *Let $U\subset H^0(W,\pi^*\mathscr{O}_{\mathbb{P}^3}(d)(-E))$ be a subspace which generates $\pi^*\mathscr{O}_{\mathbb{P}^3}(d)(-E)$. Let $0\neq\sigma\in U$. Suppose that $S=V(\sigma)$ is smooth, and that*

(a) $H^1(W,\Omega^2_W\otimes\pi^*\mathscr{O}_{\mathbb{P}^3}(d)(-E))=0$.
(b) $H^1(W,M(U)\otimes K_W\otimes\pi^*\mathscr{O}_{\mathbb{P}^3}(d)(-E))=0$.

Then the INL Theorem holds at (U,σ).

2.5. The generic tangent space is a base-point free linear system. We will prove Proposition 2.6 by applying Proposition 2.9.

Lemma 2.10. *Suppose that*

$$0=H^1(\mathbb{P}^3,\mathscr{I}_C\otimes T_{\mathbb{P}^3}(d-4))=H^1(C,T_C(d-4)). \tag{2.9}$$

Then $H^1(W,\Omega^2_W\otimes\pi^*\mathscr{O}_{\mathbb{P}^3}(d)(-E))=0$.

Proof. Since $\Omega^2_W\cong T_W\otimes K_W$ it is equivalent to prove that

$$0=H^1(W,T_W\otimes K_W\otimes\pi^*\mathscr{O}_{\mathbb{P}^3}(d)(-E))=H^1(W,T_W\otimes\pi^*\mathscr{O}_{\mathbb{P}^3}(d-4)). \tag{2.10}$$

Let $\rho\colon E\to C$ be the restriction of π. Restricting the differential of π to E, one gets an exact sequence

$$0\longrightarrow\mathscr{O}_W(E)|_E\longrightarrow\rho^*\mathscr{N}_{C/\mathbb{P}^3}\longrightarrow\xi\longrightarrow 0 \tag{2.11}$$

of sheaves on E, where ξ is an invertible sheaf. Let $\iota\colon E\hookrightarrow W$ be the inclusion map. The differential of π gives the exact sequence

$$0\longrightarrow T_W\otimes\pi^*\mathscr{O}_{\mathbb{P}^3}(d-4)\longrightarrow\pi^*T_{\mathbb{P}^3}(d-4)\longrightarrow\iota_*(\xi\otimes\rho^*\mathscr{O}_C(d-4))\longrightarrow 0. \tag{2.12}$$

Below is a piece of the associated long exact sequence of cohomology:

$$\begin{aligned}&H^0(W,\pi^*T_{\mathbb{P}^3}(d-4))\to H^0(W,\iota_*(\xi\otimes\rho^*\mathscr{O}_C(d-4)))\to\\&\to H^1(W,T_W\otimes\pi^*\mathscr{O}_{\mathbb{P}^3}(d-4))\to H^1(W,\pi^*T_{\mathbb{P}^3}(d-4)).\end{aligned} \tag{2.13}$$

We claim that $H^1(W,\pi^*T_{\mathbb{P}^3}(d-4))=0$. In fact the spectral sequence associated to π and abutting to the cohomology $H^q(W,\pi^*T_{\mathbb{P}^3}(d-4))$ gives an exact sequence

$$\begin{aligned}0\to H^1(\mathbb{P}^3,\pi_*\pi^*T_{\mathbb{P}^3}(d-4))\to H^1(W,\pi^*T_{\mathbb{P}^3}(d-4))\to&\\\to H^0(\mathbb{P}^3,R^1\pi_*\pi^*T_{\mathbb{P}^3}(d-4))\to 0.&\end{aligned} \tag{2.14}$$

Now $\pi_*\pi^*T_{\mathbb{P}^3}(d-4)\cong T_{\mathbb{P}^3}(d-4)$ and hence $H^1(\mathbb{P}^3,\pi_*\pi^*T_{\mathbb{P}^3}(d-4))=0$. Moreover $R^1\pi_*\pi^*T_{\mathbb{P}^3}(d-4)=0$ because $R^1\pi_*\mathscr{O}_W=0$, and hence

$H^1(W, \pi^* T_{\mathbb{P}^3}(d-4)) = 0$. By (2.13), in order to complete the proof it suffices to show that the map

$$H^0(W, \pi^* T_{\mathbb{P}^3}(d-4)) \to H^0(W, \iota_*(\xi \otimes \rho^* \mathscr{O}_C(d-4))) \tag{2.15}$$

is surjective. The long exact cohomology sequence associated to (2.11) gives an isomorphism

$$H^0(C, \mathscr{N}_{C/\mathbb{P}^3}(d-4)) \xrightarrow{\sim} H^0(W, \iota_*(\xi \otimes \rho^* \mathscr{O}_C(d-4))),$$

and moreover the map of (2.15) is identified with the composition

$$H^0(\mathbb{P}^3, T_{\mathbb{P}^3}(d-4)) \xrightarrow{\alpha} H^0(C, T_{\mathbb{P}^3}(d-4)|_C) \xrightarrow{\beta} H^0(C, \mathscr{N}_{C/\mathbb{P}^3}(d-4)). \tag{2.16}$$

The map α is surjective by the first vanishing in (2.9), while β is surjective by the second vanishing in (2.9). $\square$

Let $U \subset H^0(\mathbb{P}^3, \mathscr{I}_C(d))$ be a subspace which generates $\mathscr{I}_C(d)$; we let $\overline{M}(U)$ be the sheaf on $\mathbb{P}^3$ fitting into the exact sequence

$$0 \longrightarrow \overline{M}(U) \longrightarrow U \otimes \mathscr{O}_{\mathbb{P}^3} \longrightarrow \mathscr{I}_C(d) \longrightarrow 0. \tag{2.17}$$

The curve C is a local complete intersection because C is smooth, and hence $\overline{M}(U)$ is locally-free.

Lemma 2.11. *Suppose that the hypotheses of* Lemma 2.10 *hold and that in addition the sheaf $\mathscr{I}_C$ is d-regular. Let $U \subset H^0(\mathbb{P}^3, \mathscr{I}_C(d))$ be a subspace which generates $\mathscr{I}_C(d)$, and let c be its codimension. Then $\bigwedge^p \overline{M}(U)$ is $(p+c)$-regular.*

Proof. Let $\overline{M} := \overline{M}(H^0(\mathscr{I}_C(d)))$. Then $\overline{M}$ is 1-regular: in fact $H^1(\mathbb{P}^3, \overline{M}) = 0$ because the exact sequence induced by (2.17) on H^0 is exact by definition, and $H^i(\mathbb{P}^3, \overline{M}(1-i)) = 0$ for $i \geq 2$ because $\mathscr{I}_C$ is d-regular. It follows that $\bigwedge^p \overline{M}$ is p-regular (Corollary 1.8.10 of [11]). Then, arguing as in the proof of the Lemma on p. 371 of [10] (see also Example 1.8.15 of [11]) one gets that $\bigwedge^p \overline{M}(U)$ is $(p+c)$-regular. $\square$

Proof of Proposition 2.6. By hypothesis there exists a smooth point $[\sigma]$ of $(A \setminus \Delta(d))$, such that the projective tangent space $\mathbf{T}_S A$ is a base-point free codimension-1 linear subsystem of Λ. We have $\mathbf{T}_S A = \mathbb{P}(U)$, where $U \subset H^0(\pi^* \mathscr{O}_{\mathbb{P}^3}(d)(-E))$ is a codimension-1 subspace which generates $\mathscr{O}_{\mathbb{P}^3}(d)(-E)$. We will prove that the INL Theorem holds for (U, σ), and Proposition 2.6 will follow. By Joshi's Proposition 2.9 it suffices to prove that the following hold:

(a) $H^1(W, \Omega^2_W \otimes \pi^* \mathscr{O}_{\mathbb{P}^3}(d)(-E)) = 0$.
(b) $H^1(W, M(U) \otimes K_W \otimes \pi^* \mathscr{O}_{\mathbb{P}^3}(d)(-E)) = 0$.

We start by noting that, since $T_{\mathbb{P}^3}$ is -1-regular, and by hypothesis $\mathscr{I}_C$ is $(d-2)$ regular, the sheaf $\mathscr{I}_C \otimes T_{\mathbb{P}^3}$ is $(d-3)$-regular, see Proposition 1.8.9 and Remark 1.8.11 of [11]. Thus the hypotheses of Lemma 2.10 are satisfied, and hence Item (a) holds. Let us prove that Item (b) holds. Tensoring (2.8) by

$K_W \otimes \pi^*\mathscr{O}_{\mathbb{P}^3}(d)(-E) \cong \pi^*\mathscr{O}_{\mathbb{P}^3}(d-4)$ and taking cohomology we get an exact sequence

(2.18)
$$0 \to H^0(W, M(U)\otimes\pi^*\mathscr{O}_{\mathbb{P}^3}(d-4)) \to U\otimes H^0(W,\pi^*\mathscr{O}_{\mathbb{P}^3}(d-4)) \xrightarrow{\alpha} H^0(W,\pi^*\mathscr{O}_{\mathbb{P}^3}(2d-4)(-E)) \to$$
$$\to H^1(W, M(U)\otimes\pi^*\mathscr{O}_{\mathbb{P}^3}(d-4)) \to U\otimes H^1(W, K_W\otimes\pi^*\mathscr{O}_{\mathbb{P}^3}(d)(-E)).$$

Now $H^1(W, K_W \otimes \pi^*\mathscr{O}_{\mathbb{P}^3}(d)(-E)) = 0$ because by hypothesis $\pi^*\mathscr{O}_{\mathbb{P}^3}(d)(-E)$ is ample. Thus it suffices to prove that the map α is surjective. We have an identification $H^0(W, \pi^*\mathscr{O}_{\mathbb{P}^3}(d)(-E)) = H^0(\mathbb{P}^3, \mathscr{I}_C(d))$, and hence U is identified with a codimension-1 subspace of $H^0(\mathbb{P}^3, \mathscr{I}_C(d))$ that we will denote by the same symbol. Clearly it suffices to prove that the natural map

$$U \otimes H^0(\mathbb{P}^3, \mathscr{O}_{\mathbb{P}^3}(d-4)) \longrightarrow H^0(\mathbb{P}^3, \mathscr{I}_C(2d-4)) \tag{2.19}$$

is surjective. Tensorize Exact Sequence (2.17) by $\mathscr{O}_{\mathbb{P}^3}(d-4)$ and take the associated long exact sequence of cohomology: then (2.19) appears in that exact sequence, and hence it suffices to prove that $H^1(\mathbb{P}^3, \overline{M}(U)(d-4)) = 0$. By Lemma 2.11 the sheaf $\overline{M}(U)$ is 2-regular, and by hypothesis $d \geq 5$: the required vanishing follows. □

2.6. All tangent spaces at smooth points are linear systems with a base-point. We will prove Proposition 2.8. We start with an elementary result.

Lemma 2.12. *Assume that $\pi^*\mathscr{O}_{\mathbb{P}^3}(r-3)(-E)$ is very ample. Let $p \in W$ and $F \subset T_pW$ be a subspace such that one of the following holds:*

1. *$p \notin E$ and $F \neq T_pW$,*
2. *$p \notin E$ and $F = T_pW$,*
3. *$p \in E$, and $T_p(\pi^{-1}(\pi(p))) \not\subset F \subsetneq T_pE$.*

Let $X \in \Gamma_{p,F}(r)$ (see Definition 2.7*) be generic. Then X is smooth if Item (1) or (3) holds, while X has an ODP at $q = \pi(p)$ and is smooth elsewhere if Item (2) holds.*

Proof. Suppose first that (1) or (2) holds, i.e. $p \notin E$, and let $q := \pi(p)$. The linear system $\Gamma_{p,F}(r)$ has base locus $C \cup \{q\}$. In fact, let $z \in (\mathbb{P}^3 \setminus C \setminus \{q\})$; then there exists $Y \in |\mathscr{I}_C(r-2)|$ not containing z because $\pi^*\mathscr{O}_{\mathbb{P}^3}(r-2)(-E)$ is very ample, and a quadric $Q \in \Gamma_{p,F}(2)$ not containing z. Thus $Y + Q$ is an element of $\Gamma_{p,F}(r)$ which does not contain z. Hence the generic $X \in \Gamma_{p,F}(r)$ is smooth away from $C \cup \{q\}$ by Bertini. Considering $Y + Q \in \Gamma_{p,F}(r)$ as above we also get that the behaviour in q of the generic element of $\Gamma_{p,F}(r)$ is as claimed. It remains to prove that the generic $X \in \Gamma_{p,F}(r)$ is smooth at every point of C, i.e. that $\Gamma_{p,F}(r)$ is not a subset of $\Sigma(r)$. The proof that $\Gamma_{p,F}(r)$ has base locus $C \cup \{q\}$ proves also that

$$\dim \Gamma_{p,F}(r) = \dim |\mathscr{I}_C(r)| - \dim F - 1. \tag{2.20}$$

Thus in order to prove that $\Gamma_{p,F}(r)$ is not a subset of $\Sigma(r)$, it suffices to prove that for $x\in C$

$$\dim|\mathscr{I}_x^2(r)|\cap\Gamma_{p,F}(r)\le\dim|\mathscr{I}_C(r)|-\dim F-3. \tag{2.21}$$

By Item (1) of Proposition 2.3, $\dim|\mathscr{I}_x^2(r)|\cap|\mathscr{I}_C(r)|=\dim|\mathscr{I}_C(r)|-2$, and hence (2.21) is equivalent to

$$\operatorname{cod}(|\mathscr{I}_x^2(r)|\cap\Gamma_{p,F}(r),|\mathscr{I}_x^2(r)|\cap|\mathscr{I}_C(r)|)=\dim F+1. \tag{2.22}$$

We must check that imposing to $X\in|\mathscr{I}_x^2(r)|\cap|\mathscr{I}_C(r)|$ that it contains q and that $d\pi(p)(F)\subset T_qX$, gives $\dim F+1$ linearly independent conditions. By Item (1) of Proposition 2.3, there exists $Y\in|\mathscr{I}_x^2(r-2)|\cap|\mathscr{I}_C(r-2)|$ not containing q. Consider the linear subsystem $A\subset|\mathscr{I}_x^2(r)|\cap|\mathscr{I}_C(r)|$ whose elements are $Y+Q$, where $Q\in|\mathscr{O}_{\mathbb{P}^3}(2)|$; imposing to $X\in A$ that it contains q and that $d\pi(p)(F)\subset T_qX$, gives $\dim F+1$ linearly independent conditions, and hence (2.22) follows. This finishes the proof that if (1) or (2) holds, then the conclusion of the lemma holds.

Now suppose that (3) holds. Suppose that $F=\{0\}$, and let $S\in\Lambda_{p,F}(r)=|\mathscr{I}_p\otimes\pi^*\mathscr{O}_{\mathbb{P}^3}(r)(-E)|$ be generic. Then S is smooth at p because $\pi^*\mathscr{O}_{\mathbb{P}^3}(r)(-E)$ is very ample, and by Bertini's Theorem it is smooth away from p as well. In order to prove that $X=\pi(S)$ is smooth we must check that S does not contain any of the lines $\mathbf{L}_x:=\pi^{-1}(x)$ for $x\in C$. Since $\pi^*\mathscr{O}_{\mathbb{P}^3}(r)(-E)$ is very ample,

$$\operatorname{cod}(|\mathscr{I}_{\mathbf{L}_x}\otimes\pi^*\mathscr{O}_{\mathbb{P}^3}(r)(-E)|\cap|\mathscr{I}_p\otimes\pi^*\mathscr{O}_{\mathbb{P}^3}(r)(-E)|,|\mathscr{I}_p\otimes\pi^*\mathscr{O}_{\mathbb{P}^3}(r)(-E)|)=\begin{cases}1 & \text{if } x=q,\\ 2 & \text{if } x\ne q.\end{cases} \tag{2.23}$$

It follows that a generic $S\in|\mathscr{I}_p\otimes\pi^*\mathscr{O}_{\mathbb{P}^3}(r)(-E)|$ does not contain any $\mathbf{L}_x$.

We are left to deal with the case of a 1-dimensional $F\subset T_pE$ transverse to $T_p(\pi^{-1}(q))$. Let $Z\subset W$ be the 0-dimensional scheme of length 2 supported at p, with tangent space F; thus $Z\subset E$. We must prove that a generic $S\in|\mathscr{I}_Z\otimes\pi^*\mathscr{O}_{\mathbb{P}^3}(r)(-E)|$ is smooth and contains no line $\mathbf{L}_x$ where $x\in C$.

We claim that the (reduced) base-locus of $|\mathscr{I}_Z\otimes\pi^*\mathscr{O}_{\mathbb{P}^3}(r)(-E)|$ is p. In fact no $z\in(\mathbf{L}_q\setminus\{p\})$ is in the base-locus of $|\mathscr{I}_Z\otimes\pi^*\mathscr{O}_{\mathbb{P}^3}(r)(-E)|$ because $\mathbf{L}_q$ is a line and there exists $S\in|\mathscr{I}_Z\otimes\pi^*\mathscr{O}_{\mathbb{P}^3}(r)(-E)|$ which is not tangent to $\mathbf{L}_q$ at p. Moreover no $z\in(W\setminus\mathbf{L}_q)$ is in the base-locus of $|\mathscr{I}_Z\otimes\pi^*\mathscr{O}_{\mathbb{P}^3}(r)(-E)|$ because of the following argument. There exist $T\in|\mathscr{I}_p\otimes\pi^*\mathscr{O}_{\mathbb{P}^3}(r-1)(-E)|$ not containing z, and a plane $P\subset\mathbb{P}^3$ containing q and not containing $\pi(z)$; then $(T+P)\in|\mathscr{I}_Z\otimes\pi^*\mathscr{O}_{\mathbb{P}^3}(r)(-E)|$ does not contain z. This proves that the (reduced) base-locus of $|\mathscr{I}_Z\otimes\pi^*\mathscr{O}_{\mathbb{P}^3}(r)(-E)|$ is p; it follows that the generic $S\in|\mathscr{I}_Z\otimes\pi^*\mathscr{O}_{\mathbb{P}^3}(r)(-E)|$ is smooth.

We finish by showing that (2.23) holds with $\mathscr{I}_p$ replaced by $\mathscr{I}_Z$. The case $x=q$ is immediate. If $x\in C\setminus\{q\}$ we get the result by considering elements $(T+P)\in|\mathscr{I}_Z\otimes\pi^*\mathscr{O}_{\mathbb{P}^3}(r)(-E)|$ where P is a fixed plane containing q and not containing x, and $T\in|\mathscr{I}_p\otimes\pi^*\mathscr{O}_{\mathbb{P}^3}(r-1)(-E)|$. □

Remark 2.13. The proof of Lemma 2.12 shows that, if Item (2) holds, the projectivized tangent cone at q of the generic $X \in \Gamma_{p,F}(r)$ is the generic conic in $\mathbb{P}(T_q\mathbb{P}^3)$.

Proof of Proposition 2.8. Let $r \in \{d-1, d\}$. Suppose that $p \in W$, $F \subset T_pW$, and either $p \notin E$, or else $p \in E$ and (2.5) holds. By Lemma 2.12 there exists an open dense subset $\mathscr{U}_{p,F}(r) \subset \Gamma_{p,F}(r)$ such that for $X \in \mathscr{U}_{p,F}(r)$ the following holds:

(1) X is smooth if $p \notin E$ and $F \neq T_pW$, or $p \in E$.
(2) X has an ODP at $q = \pi(p)$, and is smooth elsewhere, if $p \notin E$ and $F = T_pW$.

Similary, let $j \in \{1, \ldots, n\}$, and $q \in C_j$. By Proposition 2.3 there exists an open dense subset $\mathscr{U}_{q,j}(r) \subset |\mathscr{I}_q^2(r)| \cap \Sigma_j(r)$ such that every $X \in \mathscr{U}_{q,j}(r)$ has an ODP at q and is smooth elsewhere. It will suffice to prove that if X is a very general element of $\mathscr{U}_{p,F}(r)$ or of $\mathscr{U}_{q,j}(r)$, then $\mathrm{CH}^1(X)_{\mathbb{Q}}$ is generated by $c_1(\mathscr{O}_X(1))$ and the classes of $C_1, \ldots, C_n$. Notice that if X is an element of $\mathscr{U}_{p,F}(r)$ or of $\mathscr{U}_{q,j}(r)$, then X is $\mathbb{Q}$-factorial. More precisely: if D is a Weil divisor on X then $2D$ is a Cartier divisor. Let $\mathrm{NL}(\mathscr{U}_{p,F}(d)) \subset \mathscr{U}_{p,F}(d)$ be the subset of X such that $\mathrm{Pic}(X) \otimes \mathbb{Q}$ is not generated by $\mathscr{O}_X(1)$ and $\mathscr{O}_X(2C_1), \ldots, \mathscr{O}_X(2C_n)$, and define similarly $\mathrm{NL}(\mathscr{U}_{q,j}(d)) \subset \mathscr{U}_{q,j}(d)$. Then $\mathrm{NL}(\mathscr{U}_{p,F}(d))$ is a countable union of closed subsets of $\mathscr{U}_{p,F}(d)$ (there exists a simultaneous resolution if the surfaces in $\mathscr{U}_{p,F}(d)$ are not smooth), and similarly for $\mathrm{NL}(\mathscr{U}_{q,j}(d))$. Hence it suffices to prove that $\mathscr{U}_{p,F}(d) \setminus \mathrm{NL}(\mathscr{U}_{p,F})(d)$ and $\mathscr{U}_{q,j}(d) \setminus \mathrm{NL}(\mathscr{U}_{q,j}(d))$ are not empty.

Let Y be an element of $\mathscr{U}_{p,F}(d-1)$ or of $\mathscr{U}_{q,j}(d-1)$, and let X be a generic element of $\mathscr{U}_{p,F}(d)$ or of $\mathscr{U}_{q,j}(d)$. Since $\pi^*\mathscr{O}_{\mathbb{P}^3}(d)(-E)$ is very ample and X is generic, the intersection of X and Y is reduced, and there exists an integral curve $C_0 \subset \mathbb{P}^3$ such that its irreducible decomposition is

$$X \cap Y = C_0 \cup C_1 \cup \ldots \cup C_n. \tag{2.24}$$

Now let $P \subset \mathbb{P}^3$ be a generic plane, in particular transverse to $C_0 \cup C_1 \cup \ldots \cup C_n$. Let $X = V(f)$, $Y = V(g)$ and $P = V(l)$. Let

$$\mathscr{X} := V(g \cdot l + tf) \subset \mathbb{P}^3 \times \mathbb{A}^1. \tag{2.25}$$

The projection $\mathscr{X} \to \mathbb{A}^1$ is a family of degree-d surfaces, with central fiber $Y + P$. The 3-fold $\mathscr{X}$ is singular. First $\mathscr{X}$ is singular at the points $(x, 0)$ such that $x \in X \cap Y \cap P$, and it has an ODP at each of these points because P is transverse to $C_0 \cup C_1 \cup \ldots \cup C_n$. Moreover

(I) $\mathscr{X}$ has no other singularities if we are dealing with $\mathscr{U}_{p,F}(d)$ and $F \neq T_pW$,
(II) $\mathscr{X}$ is also singular at $\{q\} \times \mathbb{A}^1$ if we are dealing with $\mathscr{U}_{p,F}(d)$ and $F = T_pW$, or if we are dealing with $\mathscr{U}_{q,j}(d)$.

We desingularize $\mathscr{X}$ as follows. The ODP's are eliminated by a small resolution (we follow p. 35 of [8], and choose a specific small resolution among the many possible ones), while to desingularize $\{q\} \times \mathbb{A}^1$ we blow-up that curve: let

$\widehat{\mathscr{X}} \to \mathscr{X}$ be the birational morphism. Then $\widehat{\mathscr{X}}$ is smooth (if $p \notin E$ and $F = T_pW$, or if we are dealing with $\mathscr{U}_{q,k}(d)$, then $\widehat{\mathscr{X}}$ is smooth over $\{q\} \times \mathbb{A}^1$ by REMARK 2.4 and REMARK 2.13).

The composition of $\widehat{\mathscr{X}} \to \mathscr{X}$ and the projection $\mathscr{X} \to \mathbb{A}^1$ is a flat family of surfaces $\varphi\colon \widehat{\mathscr{X}} \to \mathbb{A}^1$. The central fiber $\widehat{Z}_0 := \varphi^{-1}(0)$ has normal crossings, it is the union of Y and the blow-up $\widetilde{P}$ of P at the points of $X \cap Y \cap P$, the curve $Y \cap P$ being glued to its strict transform in $\widetilde{P}$. There will be an open dense $B \subset \mathbb{A}^1$ containing 0 such that $\widehat{Z}_t := \varphi^{-1}(t)$ is smooth for $t \in B \setminus \{0\}$, and it is isomorphic to $Z_t := V(g \cdot l + tf)$ in Case (I), while it is the blow-up of Z_t at q (an ODP) in Case (II). We replace $\widehat{\mathscr{X}}$ by $\varphi^{-1}(B)$ but we do not give it a new name.

One proves that if P is very general, then the following hold:

(I′) In Case (I), if t is very general in $B \setminus \{0\}$, then $\mathrm{Pic}(\widehat{Z}_t) \otimes \mathbb{Q}$ is generated by the classes of $\mathscr{O}_{\widehat{Z}_t}(1)$, $\mathscr{O}_{\widehat{Z}_t}(C_1), \ldots, \mathscr{O}_{\widehat{Z}_t}(C_n)$. (Notice that $\widehat{Z}_t = Z_t$ because we are in case (I).)

(II′) In Case (II), if t is very general in $B \setminus \{0\}$, letting $\mu_t\colon \widehat{Z}_t \to Z_t$ be the blow-up of q and $R_t \subset \widehat{Z}_t$ the exceptional curve, the group $\mathrm{Pic}(\widehat{Z}_t) \otimes \mathbb{Q}$ is generated by the classes of $\mu_t^*\mathscr{O}_{Z_t}(1)$, $\mu_t^*\mathscr{O}_{Z_t}(2C_1), \ldots, \mu_t^*\mathscr{O}_{Z_t}(2C_n)$ and $\mathscr{O}_{\widehat{Z}_t}(R_t)$.

One does this by controlling the Picard group of the degenerate fiber $\widehat{Z}_0$. As proved in [8, 12, 5] it suffices to show that the following hold:

(a) Let $\mathscr{V} \subset |\mathscr{O}_{\mathbb{P}^3}(1)|$ be the open subset of planes intersecting transversely $C_0 \cup \ldots \cup C_n$, let $I \subset (C_0 \cup \ldots \cup C_n) \times \mathscr{V}$ be the incidence subset and $\rho\colon I \to \mathscr{V}$ be the natural finite map: then the mododromy of ρ acts on a fiber $(D_0, \ldots, D_n, P)$ as the product of the symmetric groups $\mathfrak{S}_{\deg C_0} \times \ldots \times \mathfrak{S}_{\deg C_n}$.

(b) Let $j \in \{0, \ldots, n\}$, let $P \subset \mathbb{P}^3$ be a very general plane, and let $a, b \in C_j \cap P$ be distinct points; then the divisor class $a - b$ on the (smooth) curve $Y \cap P$ is not torsion.

Now Item (a) is Proposition II.2.6 of [12]. It remains to prove that (b) holds. To this end we will show that C_0 is not planar and we will control the set of planes P such that $P \cap Y$ is reducible (see the proof of Item (b) of Lemma 3.4 of [5]).

CLAIM 2.14. *The curve C_0 (see (2.24)) is not planar.*

Proof. By hypothesis $\pi^*\mathscr{O}_{\mathbb{P}^3}(d-3)(-E)$ is very ample, in particular it has a non-zero section, and hence there exists a non-zero $\tau \in H^0(\mathbb{P}^3, \mathscr{I}_C(d-3))$. Multiplying τ by sections of $\mathscr{O}_{\mathbb{P}^3}(3)$ we get that that $h^0(\mathbb{P}^3, \mathscr{I}_C(d)) \geq 20$. Now assume that C_0 is planar. Recall that $C = C_1 \cup \ldots \cup C_n$, and let

$$H^0(\mathbb{P}^3, \mathscr{I}_C(d)) \xrightarrow{\alpha} H^0(Y, \mathscr{O}_Y(d))$$

be the restriction map. Since $(C + C_0) \in |\mathscr{O}_Y(d)|$, the image of α is equal to $H^0(Y, \mathscr{O}_Y(C_0))$, and hence has dimension at most 4 because C_0 is planar.

The kernel of α has dimension 4 because Y has degree $(d-1)$. It follows that $h^0(\mathbb{P}^3, \mathscr{I}_C(d)) \leq 8$, contradicting the inequality $h^0(\mathbb{P}^3, \mathscr{I}_C(d)) \geq 20$. $\square$

Thus none of the curves $C_0, C_1, \ldots, C_n$ is planar.

LEMMA 2.15. *Let $Y \subset \mathbb{P}^3$ be a surface which is either smooth or has ODP's. The set of planes P such that $P \cap Y$ is reducible is the union of a finite set and the collection of pencils through lines of Y.*

Proof. Suppose the contrary. Then there exists a 1-dimensional family of planes P such that $P \cdot Y = C_1 + C_2$ with C_1, C_2 divisors which intersect properly, $\operatorname{supp} C_1$ is irreducible, and $\deg C_i > 1$. Next, we distinguish between the two cases:

(1) The generic P does not contain any singular point of Y.
(2) The generic P contains a single point $a \in \operatorname{sing} Y$, or two points $a, b \in \operatorname{sing} Y$.

Assume that (1) holds. Let $m_i := \deg C_i$ for $i = 1, 2$. Then

$$m_1 m_2 = (C_1 \cdot C_2)_P = (C_1 \cdot C_2)_Y = (C_1 \cdot (P - C_1))_Y = m_1 - (C_1 \cdot C_1)_Y \tag{2.26}$$

where $(C_1 \cdot C_2)_P$ is the intersection number of C_1, C_2 in the plane P, and $(C_1 \cdot C_2)_Y$ is the intersection number of C_1, C_2 in the surface Y (this makes sense because Y has ODP singularities, and hence is $\mathbb{Q}$-Cartier). The first equality of (2.26) holds by Bèzout, the second equality is proved by a local computation of the multiplicity of intersection at each point of $C_1 \cap C_2$ (one needs the hypothesis that Y is smooth at each such point). Thus (2.26) gives $(C_1 \cdot C_1)_Y = m_1(1 - m_2) < 0$, and this contradicts the hypothesis that C_1 moves in Y. If (2) holds one argues similarly. We go through the computations in the case that P contains two singular points. Let $\widetilde{\mathbb{P}}^3 \to \mathbb{P}^3$ be the blow up of $\{a, b\}$, and $\widetilde{Y}, \widetilde{P} \subset \widetilde{\mathbb{P}}^3$ be the strict transforms of Y and P respectively. By hypothesis Y has an ODP at each of its singular points and hence $\widetilde{Y}$ is smooth, and of course $\widetilde{P}$ is smooth. Let $\widetilde{C}_i$ be the strict transform of C_i in $\widetilde{\mathbb{P}}^3$. Let $r_i := \operatorname{mult}_a C_i$, $s_i := \operatorname{mult}_b C_i$. Then the equality

$$(\widetilde{C}_1 \cdot \widetilde{C}_2)_{\widetilde{P}} = (\widetilde{C}_1 \cdot \widetilde{C}_2)_{\widetilde{Y}} \tag{2.27}$$

gives

$$(\widetilde{C}_1 \cdot \widetilde{C}_2)_{\widetilde{Y}} = -(m_1 m_2 - m_1 - r_1 r_2 - s_1 s_2 + r_1 + s_1). \tag{2.28}$$

Now $r_i + s_i \leq m_i$ for $i = 1, 2$, because otherwise the line $\langle a, b \rangle$ would be contained in $Y \cap C_i$, and hence we would be considering curves residual to a line in Y, against the hypothesis. Since $r_i + s_i \leq m_i$ for $i = 1, 2$ the right-hand side of (2.28) is strictly negative, and this is a contradiction. $\square$

Now we prove that Item (b) holds. Let $j \in \{0, \ldots, n\}$. Let $a, b \in C_j$ be generic, in particular they are smooth points of Y. By LEMMA 2.15 every plane containing a, b intersects Y in an irreducible curve. Let $\widehat{Y} \to Y$ be the blow-up of the base-locus of the pencil of plane sections of Y containing a, b. Then $\widehat{Y}$ has at most A_n-singularities, and hence is $\mathbb{Q}$-factorial. Let E, F be the exceptional sets over a and b respectively, both have strictly negative self-intersection. Let $i > 0$ be such that iE and iF are Cartier. Let $\varphi\colon \widehat{Y} \to \mathbb{P}^1$ be the regular map defined by the pencil of plane sections of Y containing a, b; for $s \in \mathbb{P}^1$ we let $D_s := \varphi^{-1}(s)$. It suffices to prove that, given $r > 0$, the set of $s \in \mathbb{P}^1$ such that $\mathscr{O}_{\widehat{Y}}(riE - riF)|_{D_s}$ is trivial is finite. Assume the contrary: then, since every plane containing a, b intersects Y in an irreducible curve, there exists $\ell \in \mathbb{Q}$ such that $riE - riF \equiv \varphi^*(\ell p)$ in $\operatorname{Pic}(\widehat{Y})_{\mathbb{Q}}$, where $p \in \mathbb{P}^1$ (see the proof of Item (b) of Lemma 3.4 of [5]). It follows that the degrees of $\mathscr{O}_{\widehat{Y}}(riE - riF)$ on E and F are both equal to ℓ, and that is absurd because they have opposite signs. $\square$

3. Proof of the main result

We will prove THEOREM 0.1. Let $Q \subset \mathbb{P}^3$ be a smooth quadric and choose an isomorphism $\varphi\colon Q \xrightarrow{\sim} \mathbb{P}^1 \times \mathbb{P}^1$: we let $\mathscr{O}_Q(a, b) := \varphi^*(\mathscr{O}_{\mathbb{P}^1}(a) \boxtimes \mathscr{O}_{\mathbb{P}^1}(b))$.

PROPOSITION 3.1. *A curve in* $|\mathscr{O}_Q(2, 3)|$ *is* 3*-regular.*

Proof. Let $D \in |\mathscr{O}_Q(2, 3)|$. Considering the exact sequence $0 \to \mathscr{I}_D \to \mathscr{O}_{\mathbb{P}^3} \to \mathscr{O}_D \to 0$ we see right away that if $i = 2, 3$, then $H^i(\mathbb{P}^3, \mathscr{I}_D(3 - i)) = 0$. In order to prove that $H^1(\mathbb{P}^3, \mathscr{I}_D(2)) = 0$ we must show that $H^0(\mathbb{P}^3, \mathscr{O}_{\mathbb{P}^3}(2)) \to H^0(D, \mathscr{O}_D(2))$ is surjective. The map $H^0(\mathbb{P}^3, \mathscr{O}_{\mathbb{P}^3}(2)) \to H^0(Q, \mathscr{O}_Q(2, 2))$ is surjective, hence it suffices to prove that $H^0(Q, \mathscr{O}_Q(2, 2)) \to H^0(C, \mathscr{O}_D(2))$ is surjective. We have an exact sequence

$$0 \longrightarrow \mathscr{O}_Q(0, -1) \longrightarrow \mathscr{O}_Q(2, 2) \longrightarrow \mathscr{O}_D(2) \longrightarrow 0,$$

and since $H^1(Q, \mathscr{O}_Q(0, -1)) = 0$ the map $H^0(Q, \mathscr{O}_Q(2, 2)) \to H^0(D, \mathscr{O}_D(2))$ is indeed surjective. $\square$

Proof of THEOREM 0.1. If $d \leq 6$ there is nothing to prove, hence we may assume that $d \geq 7$. Let $n := \lfloor \frac{d-4}{3} \rfloor$. Choose disjoint smooth curves $C_1, \ldots, C_n$ such that each C_j is a $(2, 3)$-curve on a smooth quadric, and let $C := C_1 \cup \ldots \cup C_n$. We may assume that for $j \in \{1, \ldots, n\}$ the degree-0 class in $\mathrm{CH}_0(C_j)$ given by $5c_1(K_{C_j}) - 2c_1(\mathscr{O}_{C_j}(1))$ is *not* zero. Let us show that the hypotheses of THEOREM 2.1 are satisfied. Let $j \in \{1, \ldots, n\}$. We let $\pi_j : W_j \to \mathbb{P}^3$ be the blow-up of C_j, and $F_j \subset W_j$ be the exceptional divisor. Then $\pi_j^* \mathscr{O}_{\mathbb{P}^3}(3)(-F_j)$ is globally generated, and $\pi_j^* \mathscr{O}_{\mathbb{P}^3}(4)(-F_j)$ is very ample: since $d - 3 \geq 3(n-1) + 4$ it follows that $\pi^* \mathscr{O}_{\mathbb{P}^3}(d-3)(-E)$ is very ample. Let $j \in \{1, \ldots, n\}$: since $d \geq 7$ the cohomology group $H^1(C_j, T_{C_j}(d-4))$ vanishes, and hence $H^1(C, T_C(d-4)) = 0$. By PROPOSITION 3.1 and Example 1.8.32 of [11] the curve C is $3n$-regular, and since $3n \leq (d-4)$ the curve C is $(d-2)$-regular. Lastly, by construction no curve C_j is planar. We have shown that the hypotheses of THEOREM 2.1 are satisfied, and hence HYPOTHESIS 1.5 holds for $H \in |\mathscr{O}_{\mathbb{P}^3}(d)|$. Let $X \in |\mathscr{I}_C(d)|$ be smooth and very generic: since for $j \in \{1, \ldots, n\}$ the class $5c_1(K_{C_j}) - 2c_1(\mathscr{O}_{C_j}(1))$ is not zero, the decomposable classes $H^2, C_1^2, \ldots, C_n^2$ on X are linearly independent by PROPOSITION 1.7. Thus $\mathrm{DCH}_0(X)$ has rank at least $n + 1 = \lfloor \frac{d-1}{3} \rfloor$. $\square$

References

1. A. Beauville - C. Voisin, *On the Chow ring of a* $K3$ *surface*, J. Algebraic Geometry 13 (2004), pp. 417-426.
2. A. Beauville, *On the splitting of the Bloch-Beilinson filtration*, Algebraic cycles and motives. Vol. 2, London Math. Soc. Lecture Note Ser. 344, CUP, Cambridge, 2007, pp. 38-53.
3. S. Bloch - V. Srinivas, *Remarks on correspondences and algebraic cycles*, Amer. J. Math. 105 (1983), pp. 1235-1253.
4. S. Boissière - A. Sarti, *On the Néron-Severi group of surfaces with many lines*, Proc. Amer. Math. Soc. 136 (2008), pp. 3861-3867.
5. J. Brevik - S. Nollet, *Noether-Lefschetz theorem with base locus*, Int. Math. Res. Not. (2011), pp. 1220-1244.
6. L. Fu, *Decomposition of small diagonals and Chow rings of hypersurfaces and Calabi-Yau complete intersections*, Adv. Math. 244 (2013), pp. 894-924.
7. W. Fulton, *Intersection theory*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) 2. Springer-Verlag, Berlin, 1984.
8. P. Griffiths - J. Harris, *On the Noether-Lefschetz theorem and some remarks on codimension-two cycles*, Math. Ann. 271 (1985), pp. 31-51.
9. K. Joshi, *A Noether-Lefschetz Theorem and applicationss*, J. Algebraic Geom. 4 (1995), pp. 105-135.
10. S.-O. Kim, *Noether-Lefschetz locus for surfaces*, Trans. Amer. Math. Soc. 324 (1991), pp. 369-384.
11. R. Lazarsfeld, *Positivity in algebraic geometry. I. Classical setting: line bundles and linear series*, Ergebnisse der Mathematik und ihrer Grenzgebiete, 48. Springer (2004).
12. A. Lopez, *Noether-Lefschetz theory and the Picard group of projective surfaces*, Mem. Amer. Math. Soc. 89 (2011).
13. T. Shioda, *On the Picard number of a Fermat surface*, J. Fac. Sci. Univ. Tokyo Sect. IA Math. 28 (1981), pp. 725-734.
14. C. Voisin, *Théorie de Hodge et Géométrie Algébrique complexe*, SMF, Cours spécialisés 10, 2002.

Kieran G. O'Grady
Sapienza
Università di Roma

p-adic L-Functions of Automorphic Forms and Exceptional Zeros

Holger Deppe

Received: May 13, 2015

Communicated by Otmar Venjakob

Abstract. We construct p-adic L-functions for automorphic representations of GL_2 of a number field F, and show that the corresponding p-adic L-function of a modular elliptic curve E over F has an extra zero at the central point for each prime above p at which E has split multiplicative reduction, a part of the exceptional zero conjecture.

2010 Mathematics Subject Classification: 11F41, 11F67, 11F70, 11G40

Keywords and Phrases: p-adic L-function, automorphic forms, exceptional zero conjecture, Mazur-Tate-Teitelbaum conjecture

Introduction

Let F be a number field (with adele ring $\mathbb{A}_F$), and p a prime number. Let $\pi = \bigotimes_v \pi_v$ be an automorphic representation of $\mathrm{GL}_2(\mathbb{A}_F)$. Attached to π is the complex L-function $L(s,\pi)$, $s \in \mathbb{C}$, of Jacquet-Langlands [JL70]. Under certain conditions on π, we can also define a p-adic L-function $L_p(s,\pi)$ of π, with $s \in \mathbb{Z}_p$. It is related to $L(s,\pi)$ by the *interpolation property*: For every character $\chi : \mathcal{G}_p \to \mathbb{C}^*$ of finite order we have

$$L_p(0, \pi \otimes \chi) = \tau(\chi) \prod_{\mathfrak{p}|p} e(\pi_\mathfrak{p}, \chi_\mathfrak{p}) \cdot L(\tfrac{1}{2}, \pi \otimes \chi),$$

where $e(\pi_\mathfrak{p}, \chi_\mathfrak{p})$ is a certain Euler factor (see theorem 4.12 for its definition) and $\tau(\chi)$ is the Gauss sum of χ.

$L_p(s,\pi)$ was defined by Haran [Har87] in the case where π has trivial central character and $\pi_\mathfrak{p}$ is an ordinary spherical principal series representation for all $\mathfrak{p}|p$. For a totally real field F, Spieß [Sp14] has given a new construction of $L_p(s,\pi)$ that also allows for $\pi_\mathfrak{p}$ to be a special (Steinberg) representation for some $\mathfrak{p}|p$. In this article, we generalize Spieß' construction of $L_p(s,\pi)$ to

automorphic representations π of GL_2 over any number field, with arbitrary central character, and show that L_p has the conjectured number of exceptional zeros at the central point. We assume that π is ordinary at all primes $\mathfrak{p}|p$ (cf. definition 2.3), that π_v is discrete of weight 2 at all real infinite places v, and is the principal series representation $\sigma(|\cdot|^{1/2}, |\cdot|^{-1/2})$ at the complex places.
We define a p-adic measure μ_π, which heuristically is the image under the global reciprocity map of a product of certain local distributions $\mu_{\pi_\mathfrak{p}}$ on $F_\mathfrak{p}^*$ attached to $\pi_\mathfrak{p}$ for $\mathfrak{p}|p$ and a Whittaker function times the Haar measure on the group of p-ideles $\mathbb{I}^p = \prod'_{v\nmid p} F_v^*$.
Then we can define the p-adic L-function of π as an integral with respect to μ_π over the Galois group $\mathcal{G}_p$ of the maximal abelian extension that is unramified outside p and ∞; it is naturally a t-variable function, where t is the $\mathbb{Z}_p$-rank of $\mathcal{G}_p$:

$$L_p(\underline{s}, \pi) := L_p(s_1, \ldots, s_t, \pi) := \int_{\mathcal{G}_p} \prod_{i=1}^{t} \exp_p(s_i \ell_i(\gamma)) \mu_\pi(d\gamma)$$

for $s_1, \ldots, s_t \in \mathbb{Z}_p$, where the ℓ_i are $\mathbb{Z}_p$-valued homomorphisms corresponding to the t independent $\mathbb{Z}_p$-extensions of F (cf. section 4.7 for their definition).
For a modular elliptic curve E over F corresponding to π (i.e. the local L-factors of the Hasse-Weil L-function $L(E,s)$ and of the automorphic L-function $L(s-\frac{1}{2}, \pi)$ coincide at all places v of F), our construction allows us to define the p-adic L-function of E as $L_p(E, \underline{s}) := L_p(\underline{s}, \pi)$. The condition that π be ordinary at all $\mathfrak{p}|p$ means that E must have good ordinary or multiplicative reduction at all places $\mathfrak{p}|p$ of F.
The *exceptional zero conjecture* (formulated by Mazur, Tate and Teitelbaum [MTT86] for $F = \mathbb{Q}$, and by Hida [Hi09] for totally real F) states that

$$\mathrm{ord}_{s=0} L_p(E, s) \geq n, \tag{1}$$

where n is the number of $\mathfrak{p}|p$ at which E has split multiplicative reduction, and gives an explicit formula for the value of the n-th derivative $L_p^{(n)}(E, 0)$ as a multiple of certain L-invariants times $L(E, 1)$. The conjecture was proved in the case $F = \mathbb{Q}$ by Greenberg and Stevens [GS93] and independently by Kato, Kurihara and Tsuji, and for totally real fields F by Spieß [Sp14]. In this article, we prove (1) for all number fields F.

The structure of this article is as follows: In chapter 2, we describe the local distributions $\mu_{\pi_\mathfrak{p}}$ on $F_\mathfrak{p}^*$; they are the image of a Whittaker functional under a map δ on the dual of $\pi_\mathfrak{p}$. For constructing δ, we describe $\pi_\mathfrak{p}$ in terms of what we call the "Bruhat-Tits graph" of $F_\mathfrak{p}^2$: the directed graph whose vertices (resp. edges) are the lattices of $F_\mathfrak{p}^2$ (resp. inclusions between lattices). Roughly speaking, it is a covering of the (directed) Bruhat-Tits tree of $\mathrm{GL}_2(F_\mathfrak{p})$ with fibres $\cong \mathbb{Z}$. When $\pi_\mathfrak{p}$ is the Steinberg representation, $\mu_\mathfrak{p}$ can actually be extended to all of $F_\mathfrak{p}$.
In chapter 3, we attach a p-adic distribution μ_ϕ to any map $\phi(U, x^p)$ of an open compact subset $U \subseteq F_p^* := \prod_{\mathfrak{p}|p} F_\mathfrak{p}^*$ and an idele $x^p \in \mathbb{I}^p$ (satisfying certain

conditions). Integrating ϕ over all the infinite places, we get a cohomology class $\kappa_\phi \in H^d(F^{*\prime}, \mathcal{D}_f(\mathbb{C}))$ (where $d = r + s - 1$ is the rank of the group of units of F, $F^{*\prime} \cong F^*/\mu_F$ is a maximal torsion-free subgroup of F^*, and $\mathcal{D}_f(\mathbb{C})$ is a space of distributions on the finite ideles of F). We show that μ_ϕ can be described solely in terms of κ_ϕ, and μ_ϕ is a (vector-valued) p-adic measure if κ_ϕ is "integral", i.e. if it lies in the image of $H^d(F^{*\prime}, \mathcal{D}_f(R))$, for a Dedekind ring R consisting of "p-adic integers".
In chapter 4, we define a map ϕ_π by

$$\phi_\pi(U, x^p) := \sum_{\zeta \in F^*} \mu_{\pi_p}(\zeta U) W^p \begin{pmatrix} \zeta x^p & 0 \\ 0 & 1 \end{pmatrix}$$

($U \subseteq F_p^*$ compact open, $x^p \in \mathbb{I}^p$). ϕ_π satisfies the conditions of chapter 3, and we show that $\kappa_\pi := \kappa_{\phi_\pi}$ is integral by "lifting" the map $\phi_\pi \mapsto \kappa_\pi$ to a function mapping an automorphic form to a cohomology class in $H^d(\mathrm{GL}_2(F)^+, \mathcal{A}_f)$, for a certain space of functions $\mathcal{A}_f$. (Here $\mathrm{GL}_2(F)^+$ is the subgroup of $M \in \mathrm{GL}_2(F)$ with totally positive determinant.) For this, we associate to each automorphic form φ a harmonic form ω_φ on a generalized upper-half space $\mathcal{H}_\infty$, which we can integrate between any two cusps in $\mathbb{P}^1(F)$.
Then we can define the p-adic L-function $L_p(\underline{s}, \pi) := L_p(\underline{s}, \kappa_\pi)$ as above, with $\mu_\pi := \mu_{\phi_\pi}$. By a result of Harder [Ha87], $H^d(\mathrm{GL}_2(F)^+, \mathcal{A}_f)_\pi$ is one-dimensional, which implies that $L_p(\underline{s}, \pi)$ has values in a one-dimensional $\mathbb{C}_p$-vector space. Finally, we formulate an exceptional zero conjecture (conjecture 4.15) for all number fields F, and show that $L_p(\underline{s}, \pi)$ satisfies (1).

Acknowledgements

This paper is based on my Ph.D. thesis "p-adic L-functions of automorphic forms" [De13], submitted at Bielefeld University in August 2013.
I would like to thank Michael Spieß for suggesting and advising the thesis, and for many helpful discussions. I am also thankful to Werner Hoffmann for a useful discussion, to the referee for pointing out a few mistakes in an earlier version of this paper, and to the CRC 701, 'Spectral Structures and Topological Methods in Mathematics', for providing financial support during most of my studies.

Contents

1 Preliminaries

Let $\mathcal{X}$ be a totally disconnected locally compact topological space, R a topological Hausdorff ring. We denote by $C(\mathcal{X}, R)$ the ring of continuous maps $\mathcal{X} \to R$, and let $C_c(\mathcal{X}, R) \subseteq C(\mathcal{X}, R)$ be the subring of compactly supported maps. When R has the discrete topology, we also write $C^0(\mathcal{X}, R) := C(\mathcal{X}, R)$, $C^0_c(\mathcal{X}, R) := C_c(\mathcal{X}, R)$.
We denote by $\mathfrak{Co}(\mathcal{X})$ the set of all compact open subsets of $\mathcal{X}$, and for an R-module M we denote by $\mathrm{Dist}(\mathcal{X}, M)$ the R-module of M-valued distributions on $\mathcal{X}$, i.e. the set of maps $\mu : \mathfrak{Co}(\mathcal{X}) \to M$ such that $\mu(\bigcup_{i=1}^n U_i) = \sum_{i=1}^n \mu(U_i)$ for any pairwise disjoint sets $U_i \in \mathfrak{Co}(\mathcal{X})$.
For an open set $H \subseteq \mathcal{X}$, we let $1_H \in C(\mathcal{X}, R)$ be the R-valued indicator function of H on $\mathcal{X}$.
Throughout this paper, we fix a prime p and embeddings $\iota_\infty : \overline{\mathbb{Q}} \hookrightarrow \mathbb{C}$, $\iota_p : \overline{\mathbb{Q}} \hookrightarrow \mathbb{C}_p$. Let $\overline{\mathcal{O}}$ denote the valuation ring of $\overline{\mathbb{Q}}$ with respect to the p-adic valuation induced by ι_p.

We write $G := \mathrm{GL}_2$ throughout the article, and let B denote the Borel subgroup of upper triangular matrices, T the maximal torus (consisting of all diagonal matrices), and Z the center of G.
For a number field F, we let $G(F)^+ \subseteq G(F)$ and $B(F)^+ \subseteq B(F)$ denote the corresponding subgroups of matrices with totally positive determinant, i.e. $\sigma(\det(g))$ is positive for each real embedding $\sigma : F \hookrightarrow \mathbb{R}$. (If F is totally complex, this is an empty condition, so we have $G(F)^+ = G(F)$, $B(F)^+ = B(F)$ in this case.) Similarly, we define $G(\mathbb{R})^+$ and $G(\mathbb{C})^+ = G(\mathbb{C})$.

1.1 p-ADIC MEASURES

DEFINITION 1.1. Let $\mathcal{X}$ be a compact totally disconnected topological space. For a distribution $\mu : \mathfrak{Co}(\mathcal{X}) \to \mathbb{C}$, consider the extension of μ to the $\mathbb{C}_p$-linear map $C^0(\mathcal{X}, \mathbb{C}_p) \to \mathbb{C}_p \otimes_{\overline{\mathbb{Q}}} \mathbb{C}$, $f \mapsto \int f d\mu$. If its image is a finitely-generated $\mathbb{C}_p$-vector space, μ is called a *p-adic measure.*

We denote the space of p-adic measures on $\mathcal{X}$ by $\mathrm{Dist}^b(\mathcal{X}, \mathbb{C}) \subseteq \mathrm{Dist}(\mathcal{X}, \mathbb{C})$. It is easily seen that μ is a p-adic measure if and only if the image of μ, considered as a map $C^0(\mathcal{X}, \mathbb{Z}) \to \mathbb{C}$, is contained in a finitely generated $\overline{\mathcal{O}}$-module. A p-adic measure can be integrated against any continuous function $f \in C(\mathcal{X}, \mathbb{C}_p)$.

2 LOCAL RESULTS

For this chapter, let F be a finite extension of $\mathbb{Q}_p$, $\mathcal{O}_F$ its ring of integers, ϖ its uniformizer and $\mathfrak{p} = (\varpi)$ the maximal ideal. Let q be the cardinality of $\mathcal{O}_F/\mathfrak{p}$, and set $U := U^{(0)} := \mathcal{O}_F^\times$, $U^{(n)} := 1 + \mathfrak{p}^n \subseteq U$ for $n \geq 1$.
We fix an additive character $\psi : F \to \overline{\mathbb{Q}}^*$ with $\ker \psi \supseteq \mathcal{O}_F$ and $\mathfrak{p}^{-1} \not\subseteq \ker \psi$. [1]
We let $|\cdot|$ be the absolute value on F^* (normalized by $|\varpi| = q^{-1}$), $\mathrm{ord} = \mathrm{ord}_\varpi$ the additive valuation, and dx the Haar measure on F normalized by $\int_{\mathcal{O}_F} dx = 1$. We define a (Haar) measure on F^* by $d^\times x := \frac{q}{q-1}\frac{dx}{|x|}$ (so $\int_{\mathcal{O}_F^\times} d^\times x = 1$).

2.1 GAUSS SUMS

Recall that the *conductor* of a character $\chi : F^* \to \mathbb{C}^*$ is by definition the largest ideal $\mathfrak{p}^n$, $n \geq 0$, such that $\ker \chi \supseteq U^{(n)}$, and that χ is *unramified* if its conductor is $\mathfrak{p}^0 = \mathcal{O}_F$.

DEFINITION 2.1. Let $\chi : F^* \to \mathbb{C}^*$ be a quasi-character with conductor $\mathfrak{p}^f$. The *Gauss sum* of χ (with respect to ψ) is defined by

$$\tau(\chi) := [U : U^{(f)}] \int_{\varpi^{-f}U} \psi(x)\chi(x) d^\times x.$$

[1] Note that there is in general no ψ such that $\ker(\psi) = \mathcal{O}_F$, since $\mathfrak{p}^{-1}/\mathcal{O}_F$ has more than p points of order p if $F|\mathbb{Q}_p$ has inertia index > 1.

For a locally constant function $g : F^* \to \mathbb{C}$, we define

$$\int_{F^*} g(x)dx := \lim_{n\to\infty} \int_{x\in F^*, -n\leq \mathrm{ord}(x)\leq n} g(x)dx,$$

whenever that limit exists.

Lemma 2.2. *Let $\chi : F^* \to \mathbb{C}^*$ be a quasi-character with conductor $\mathfrak{p}^f$. For $f = 0$, assume $|\chi(\varpi)| < q$. Then we have*

$$\int_{F^*} \chi(x)\psi(x)dx = \begin{cases} \frac{1-\chi(\varpi)^{-1}}{1-\chi(\varpi)q^{-1}} & \text{if } f = 0 \\ \tau(\chi) & \text{if } f > 0. \end{cases}$$

(Cf. [Sp14], lemma 3.4.)

2.2 Tamely ramified representations of $\mathrm{GL}_2(F)$

For an ideal $\mathfrak{a} \subset \mathcal{O}_F$, let $K_0(\mathfrak{a}) \subseteq G(\mathcal{O}_F)$ be the subgroup of matrices congruent to an upper triangular matrix modulo $\mathfrak{a}$.
Let $\pi : \mathrm{GL}_2(F) \to \mathrm{GL}(V)$ be an irreducible admissible infinite-dimensional representation on a $\mathbb{C}$-vector space V, with central quasicharacter χ. It is well-known (e.g [Ge75], Thm. 4.24) that there exists a maximal ideal $\mathfrak{c}(\pi) = \mathfrak{c} \subset \mathcal{O}_F$, the *conductor* of π, such that the space $V^{K_0(\mathfrak{c}),\chi} = \{v \in V | \pi(g)v = \chi(a)v \; \forall g = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in K_0(\mathfrak{c})\}$ is non-zero (and in fact one-dimensional). A representation π is called *tamely ramified* if its conductor divides $\mathfrak{p}$.
If π is tamely ramified, then π is the spherical resp. special representation $\pi(\chi_1, \chi_2)$ (in the notation of [Ge75] or [Sp14]):
If the conductor is $\mathcal{O}_F$, π is (by definition) spherical and thus a principal series representation $\pi(\chi_1, \chi_2)$ for two unramified quasi-characters χ_1 and χ_2 with $\chi_1\chi_2^{-1} \neq |\cdot|^{\pm 1}$ ([Bu98], Thm. 4.6.4).
If the conductor is $\mathfrak{p}$, then $\pi = \pi(\chi_1, \chi_2)$ with $\chi_1\chi_2^{-1} = |\cdot|^{\pm 1}$.
For $\alpha \in \mathbb{C}^*$, we define a character $\chi_\alpha : F^* \to \mathbb{C}^*$ by $\chi_\alpha(x) := \alpha^{\mathrm{ord}(x)}$.
So let now $\pi = \pi(\chi_1, \chi_2)$ be a tamely ramified irreducible admissible infinite-dimensional representation of $\mathrm{GL}_2(F)$; in the special case, we assume χ_1 and χ_2 to be ordered such that $\chi_1 = |\cdot|\chi_2$.
Set $\alpha_i := \chi_i(\varpi)\sqrt{q} \in \mathbb{C}^*$ for $i = 1, 2$. (We also write $\pi = \pi_{\alpha_1,\alpha_2}$ sometimes.) Set $a := \alpha_1 + \alpha_2$, $\nu := \alpha_1\alpha_2/q$. Define a distribution $\mu_{\alpha_1,\nu} := \mu_{\alpha_1/\nu} := \psi(x)\chi_{\alpha_1/\nu}(x)dx$ on F^*.
For later use, we will need the following condition on the α_i:

Definition 2.3. Let $\pi = \pi_{\alpha_1,\alpha_2}$ be tamely ramified. π is called *ordinary* if a and ν both lie in $\overline{\mathcal{O}}^*$ (i.e. they are p-adic units in $\overline{\mathbb{Q}}$). Equivalently, this means that either $\alpha_1 \in \overline{\mathcal{O}}^*$ and $\alpha_2 \in q\overline{\mathcal{O}}^*$, or vice versa.

Proposition 2.4. *Let $\chi : F^* \to \mathbb{C}^*$ be a quasi-character with conductor $\mathfrak{p}^f$; for $f = 0$, assume $|\chi(\varpi)| < |\alpha_2|$. Then the integral $\int_{F^*} \chi(x)\mu_{\alpha_1/\nu}(dx)$ converges* *and we have*

$$\int_{F^*} \chi(x)\mu_{\alpha_1/\nu}(dx) = e(\alpha_1, \alpha_2, \chi)\tau(\chi)L(\tfrac{1}{2}, \pi \otimes \chi),$$

where

$$e(\alpha_1,\alpha_2,\chi) = \begin{cases} \frac{(1-\alpha_1\chi(\varpi)q^{-1})(1-\alpha_2\chi(\varpi)^{-1}q^{-1})(1-\alpha_2\chi(\varpi)q^{-1})}{(1-\chi(\varpi)\alpha_2^{-1})}, & f=0 \text{ and } \pi \text{ spherical,} \\ \frac{(1-\alpha_1\chi(\varpi)q^{-1})(1-\alpha_2\chi(\varpi)^{-1}q^{-1})}{(1-\chi(\varpi)\alpha_2^{-1})}, & f=0 \text{ and } \pi \text{ special,} \\ \left(\frac{\alpha_1}{\nu}\right)^{-f} = \left(\frac{\alpha_2}{q}\right)^{f}, & f>0, \end{cases}$$

and where we assume the right-hand side to be continuously extended to the potential removable singularities at $\chi(\varpi) = q/\alpha_1$ or $= q/\alpha_2$.

Proof. This follows immediately from lemma 2.2 and the definition of the (Jacquet-Langlands) L-function. □

2.3 The Bruhat-Tits graph

Let $\tilde{\mathcal{V}}$ denote the set of lattices (i.e. submodules isomorphic to $\mathcal{O}_F^2$) in F^2, and let $\tilde{\mathcal{E}}$ be the set of all inclusion maps between two lattices; for such a map $e : v_1 \hookrightarrow v_2$ in $\tilde{\mathcal{E}}$, we define $o(e) := v_1$, $t(e) := v_2$. Then the pair $\tilde{\mathcal{T}} := (\tilde{\mathcal{V}}, \tilde{\mathcal{E}})$ is naturally a directed graph, connected, with no directed cycles (specifically, $\tilde{\mathcal{E}}$ induces a partial ordering on $\tilde{\mathcal{V}}$). For each $v \in \tilde{\mathcal{V}}$, there are exactly $q+1$ edges beginning (resp. ending) in v, each.

Recall that the *Bruhat-Tits tree* $\mathcal{T} = (\mathcal{V}, \vec{\mathcal{E}})$ of $G(F)$ is the directed graph whose vertices are homothety classes of lattices of F^2 (i.e. $\mathcal{V} = \tilde{\mathcal{V}}/\sim$, where $v \sim \varpi^i v$ for all $i \in \mathbb{Z}$), and the directed edges $\overline{e} \in \vec{\mathcal{E}}$ are homothety classes of inclusions of lattices. We can define maps $o, t : \vec{\mathcal{E}} \to \mathcal{V}$ analogously. For each edge $\overline{e} \in \vec{\mathcal{E}}$, there is an opposite edge $\overline{e}' \in \vec{\mathcal{E}}$ with $o(\overline{e}') = t(\overline{e})$, $t(\overline{e}') = o(\overline{e})$; and the undirected graph underlying $\mathcal{T}$ is simply connected. We have a natural "projection map" $\pi : \tilde{\mathcal{T}} \to \mathcal{T}$, mapping each lattice and each homomorphism to its homothety class. Choosing a (set-theoretic) section $s : \mathcal{V} \to \tilde{\mathcal{V}}$, we get a bijection $\mathcal{V} \times \mathbb{Z} \xrightarrow{\cong} \tilde{\mathcal{V}}$ via $(v, i) \mapsto \varpi^i s(v)$.

The group $G(F)$ operates on $\tilde{\mathcal{V}}$ via its standard action on F^2, i.e. $gv = \{gx | x \in v\}$ for $g \in G(F)$, and on $\tilde{\mathcal{E}}$ by mapping $e : v_1 \to v_2$ to the inclusion map $ge : gv_1 \to gv_2$. The stabilizer of the standard vertex $v_0 := \mathcal{O}_F^2$ is $G(\mathcal{O}_F)$.

For a directed edge $\overline{e} \in \vec{\mathcal{E}}$ of the Bruhat-Tits tree $\mathcal{T}$, we define $U(\overline{e})$ to be the set of ends of $\overline{e}$ (cf. [Se80]/[Sp14]); it is a compact open subset of $\mathbb{P}^1(F)$, and we have $gU(\overline{e}) = U(g\overline{e})$ for all $g \in G(F)$.

For $n \in \mathbb{Z}$, we set $v_n := \mathcal{O}_F \oplus \mathfrak{p}^n \in \tilde{\mathcal{V}}$, and denote by e_n the edge from v_{n+1} to v_n; the "decreasing" sequence $(\pi(e_{-n}))_{n \in \mathbb{Z}}$ is the geodesic from ∞ to 0. (The geodesic from 0 to ∞ traverses the $\pi(v_n)$ in the natural order of $n \in \mathbb{Z}$.) We have $U(\pi(e_n)) = \mathfrak{p}^{-n}$ for each n.

On $\mathcal{T}$, we have the height function $h : \mathcal{V} \to \mathbb{Z}$ (cf. [BL95]) defined as follows: The geodesic ray from $v \in \mathcal{V}$ to ∞ must contain some $\pi(v_n)$ ($n \in \mathbb{Z}$), since it has non-empty intersection with $A := \{\pi(v_n)|n \in \mathbb{Z}\}$; we define $h(v) := n - d(v, \pi(v_n))$ for any such v_n. This is easily seen to be well-defined, and satisfies $h(\pi(v_n)) = n$ for all $n \in \mathbb{Z}$. We have the following lemma:

LEMMA 2.5. *(a) For all $\overline{e} \in \mathcal{E}$, we have*

$$h(t(\overline{e})) = \begin{cases} h(o(\overline{e})) + 1 & \text{if } \infty \in U(\overline{e}), \\ h(o(\overline{e})) - 1 & \text{otherwise.} \end{cases}$$

(b) For $a \in F^*$, $b \in F$, $\overline{v} \in \mathcal{V}$ we have

$$h\left(\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \overline{v}\right) = h(\overline{v}) - \operatorname{ord}_{\varpi}(a).$$

(Cf. [Sp14], lemma 3.6)
Let R be a ring, M an R-module. We let $C(\tilde{\mathcal{V}}, M)$ be the R-module of maps $\phi : \tilde{\mathcal{V}} \to M$, and $C(\tilde{\mathcal{E}}, M)$ the R-module of maps $\tilde{\mathcal{E}} \to M$. Both are $G(F)$-modules via $(g\phi)(v) := \phi(g^{-1}v)$, $(gc)(e) := c(g^{-1}e)$.
We let $\mathcal{C}_c(\tilde{\mathcal{V}}, M) \subseteq C(\tilde{\mathcal{V}}, M)$ and $\mathcal{C}_c(\tilde{\mathcal{E}}, M) \subseteq C(\tilde{\mathcal{E}}, M)$ be the ($G(F)$-stable) submodules of maps with compact support, i.e. maps that are zero outside a finite set. We get pairings

$$\langle \text{-}\,,\text{-} \rangle : C_c(\tilde{\mathcal{V}}, R) \times C(\tilde{\mathcal{V}}, M) \to M, \quad \langle \phi_1, \phi_2 \rangle := \sum_{v \in \tilde{\mathcal{V}}} \phi_1(v)\phi_2(v) \tag{2}$$

and

$$\langle \text{-}\,,\text{-} \rangle : C_c(\tilde{\mathcal{E}}, R) \times C(\tilde{\mathcal{E}}, M) \to M, \quad \langle c_1, c_2 \rangle := \sum_{e \in \tilde{\mathcal{E}}} c_1(v)c_2(v). \tag{3}$$

We define Hecke operators $T, N : \mathcal{C}(\tilde{\mathcal{V}}, M) \to \mathcal{C}(\tilde{\mathcal{V}}, M)$ by

$$T\phi(v) = \sum_{t(e)=v} \phi(o(e)) \quad \text{and} \quad N\phi := \varpi\phi \text{ (i.e. } N\phi(v) = \phi(\varpi^{-1}v))$$

for all $v \in \tilde{\mathcal{V}}$. These restrict to operators on $C_c(\tilde{\mathcal{V}}, R)$, which we sometimes denote by T_c and N_c for emphasis. With respect to (2), T_c is adjoint to TN, and N_c is adjoint to its inverse operator $N^{-1} : \mathcal{C}_c(\tilde{\mathcal{V}}, R) \to C_c(\tilde{\mathcal{V}}, R)$.

T and N obviously commute, and we have the following Hecke structure theorem on compactly supported functions on $\tilde{\mathcal{V}}$ (an analogue of [BL95], Thm. 10):

THEOREM 2.6. *$C_c(\tilde{\mathcal{V}}, R)$ is a free $R[T, N^{\pm 1}]$-module (where $R[T, N^{\pm 1}]$ is the ring of Laurent polynomials in N over the polynomial ring $R[T]$, with N and T commuting).*

Proof. Fix a vertex $v_0 \in \tilde{\mathcal{V}}$. For each $n \geq 0$, let C_n be the set of vertices $v \in \tilde{\mathcal{V}}$ such that there is a directed path of length n from v_0 to v in $\tilde{\mathcal{V}}$, and such that $d(\pi(v_0), \pi(v)) = n$ in the Bruhat-Tits tree $\mathcal{T}$. So $C_0 = \{v_0\}$, and C_n is a lift of the "circle of radius n around v_0" in $\mathcal{T}$, in the parlance of [BL95].
One easily sees that $\bigcup_{n=0}^{\infty} C_n$ is a complete set of representatives for the projection map $\pi : \tilde{\mathcal{V}} \to \mathcal{V}$; specifically, for $n > 1$ and a given $v \in C_{n-1}$, C_n contains exactly q elements adjacent to v in $\tilde{\mathcal{V}}$; and we can write $\tilde{\mathcal{V}}$ as a disjoint union $\bigcup_{j\in\mathbb{Z}} \bigcup_{n=0}^{\infty} N^j(C_n)$.
We further define $V_0 := \{v_0\}$ and choose subsets $V_n \subseteq C_n$ as follows: We let V_1 be any subset of cardinality q. For $n > 1$, we choose $q-1$ out of the q elements of C_n adjacent to v', for every $v' \in C_{n-1}$, and let V_n be the union of these elements for all $v' \in C_{n-1}$. Finally, we set

$$H_{n,j} := \{\phi \in C_c(\tilde{\mathcal{V}}, R) |\, \mathrm{Supp}(\phi) \subseteq \bigcup_{i=0}^{n} N^j(C_i)\} \quad \text{for each } n \geq 0, j \in \mathbb{Z},$$

$H_n := \bigcup_{j\in\mathbb{Z}} H_{n,j}$, and $H_{-1} := H_{-1,j} := \{0\}$. (For ease of notation, we identify $v \in \tilde{\mathcal{V}}$ with its indicator function $1_{\{v\}} \in C_c(\tilde{\mathcal{V}}, R)$ in this proof.)
Define $T' : C_c(\tilde{\mathcal{V}}, R) \to C_c(\tilde{\mathcal{V}}, R)$ by

$$T'(\phi)(v) := \sum_{\substack{t(e)=(v),\\ o(e)\in N^j(C_n)}} \phi(o(e)) \quad \text{for each } v \in N^j(C_{n-1}), j \in \mathbb{Z};$$

T' can be seen as the "restriction to one layer" $\bigcup_{n=0}^{\infty} N^j(C_n)$ of T. We have $T'(v) \equiv T(v) \mod H_{n-1}$ for each $v \in H_n$, since the "missing summand" of T' lies in H_{n-1}.
We claim that for each $n \geq 0$, the set $X_{n,j} := \bigcup_{i=0}^{n} N^j T^{n-i}(V_i)$ is an R-basis for $H_{n,j}/H_{n-1,j}$. By the above congruence, we can replace T by T' in the definition of $X_{n,j}$.
The claim is clear for $n = 0$. So let $n \geq 1$, and assume the claim to be true for all $n' \leq n$. For each $v \in C_{n-1}$, the q points in C_n adjacent to v are generated by the $q-1$ of these points lying in V_n, plus $T'v$ (which just sums up these q points). By induction hypothesis, v is generated by $X_{n-1,0}$, and thus (taking the union over all v), C_n is generated by $T'(X_{n-1,0}) \cup V_n = X_{n,0}$. Since the cardinality of $X_{n,0}$ equals the R-rank of $H_{n,0}/H_{n-1,0}$ (both are equal to $(q+1)q^{n-1}$), $X_{n,0}$ is in fact an R-basis.
Analoguously, we see that $H_{n,j}/H_{n-1,j}$ has $N^j(X_{n,0}) = X_{n,j}$ as a basis, for each $j \in \mathbb{Z}$.
From the claim, it follows that $\bigcup_{j\in\mathbb{Z}} X_{n,j}$ is an R-basis of H_n/H_{n-1} for each n, and that $V := \bigcup_{n=0}^{\infty} V_n$ is an $R[T, N^{\pm 1}]$-basis of $C_c(\tilde{\mathcal{V}}, R)$.

□

For $a \in R$ and $\nu \in R^*$, we let $\tilde{\mathcal{B}}_{a,\nu}(F, R)$ be the "common cokernel" of $T - a$ and $N - \nu$ in $C_c(\tilde{\mathcal{V}}, R)$, namely $\tilde{\mathcal{B}}_{a,\nu}(F, R) := C_c(\tilde{\mathcal{V}}, R)/(\mathrm{Im}(T-a) + \mathrm{Im}(N-\nu))$;

dually, we define $\tilde{\mathcal{B}}^{a,\nu}(F,M) := \ker(T-a)\cap\ker(N-\nu) \subseteq C(\tilde{\mathcal{V}},M)$.

For a lattice $v \in \tilde{\mathcal{V}}$, we define a valuation ord_v on F as follows: For $w \in F^2$, the set $\{x \in F | xw \in v\}$ is some fractional ideal $\varpi^m\mathcal{O}_F \subseteq F$ ($m \in \mathbb{Z}$); we set $\operatorname{ord}_v(w) := m$. This map can also be given explicitly as follows: Let λ_1, λ_2 be a basis of v. We can write any $w \in F^2$ as $w = x_1\lambda_1 + x_2\lambda_2$; then we have $\operatorname{ord}_v(w) = \min\{\operatorname{ord}_\varpi(x_1), \operatorname{ord}_\varpi(x_2)\}$. This gives a "valuation" map on F^2, as one easily checks. We restrict it to $F \cong F\times\{0\} \hookrightarrow F^2$ to get a valuation ord_v on F, and consider especially the value at $e_1 := (1,0)$.

LEMMA 2.7. *Let $\alpha,\nu \in R^*$, and put $a := \alpha + q\nu/\alpha$. Define a map $\varrho = \varrho_{\alpha,\nu} : \tilde{\mathcal{V}} \to R$ by $\varrho(v) := \alpha^{h(\pi(v))}\nu^{-\operatorname{ord}_v(e_1)}$. Then $\varrho \in \tilde{\mathcal{B}}^{a,\nu}(F,R)$.*

Proof. One easily sees that $\left(v \mapsto \nu^{-\operatorname{ord}_v(e_1)}\right) \in \ker(N-\nu)$. It remains to show that $\varrho \in \ker(T-a)$:
We have the Iwasawa decomposition $G(F) = B(F)G(\mathcal{O}_F) = \{\left(\begin{smallmatrix} * & * \\ 0 & 1\end{smallmatrix}\right)\}Z(F)G(\mathcal{O}_F)$; thus every vertex in $\tilde{\mathcal{V}}$ can be written as $\varpi^i v$ with $v = \left(\begin{smallmatrix} a & b \\ 0 & 1\end{smallmatrix}\right)v_0$, with $i \in \mathbb{Z}$, $a \in F^*, b \in F$.
Now the lattice $v = \left(\begin{smallmatrix} a & b \\ 0 & 1\end{smallmatrix}\right)v_0$ is generated by the vectors $\lambda_1 = \left(\begin{smallmatrix} a \\ 0\end{smallmatrix}\right)$ and $\lambda_2 = \left(\begin{smallmatrix} b \\ 1\end{smallmatrix}\right) \in \mathcal{O}_F^2$, so $e_1 = a^{-1}\lambda_1$ and thus $\operatorname{ord}_v(e_1) = \operatorname{ord}_\varpi(a^{-1}) = -\operatorname{ord}_\varpi(a)$. The $q+1$ neighbouring vertices v' for which there exists an $e \in \tilde{\mathcal{E}}$ with $o(e) = v', t(e) = v$ are given by $N_i v$, $i \in \{\infty\}\cup\mathcal{O}_F/\mathfrak{p}$, with $N_\infty := \left(\begin{smallmatrix} 1 & 0 \\ 0 & \varpi\end{smallmatrix}\right)$, and $N_i := \left(\begin{smallmatrix} \varpi & i \\ 0 & 1\end{smallmatrix}\right)$ where $i \in \mathcal{O}_F$ runs through a complete set of representatives mod ϖ. By lemma 2.5, $h(\pi(N_\infty v)) = h(\pi(v))+1$ and $h(\pi(N_i v)) = h(\pi(v))-1$ for $i \neq \infty$. By considering the basis $\{N_i\lambda_1, N_i\lambda_2\}$ of $N_i v$ for each N_i, we see that $\operatorname{ord}_{N_\infty v}(e_1) = \operatorname{ord}_v(e_1)$ and $\operatorname{ord}_{N_i v}(e_1) = \operatorname{ord}_v(e_1) - 1$ for $i \neq \infty$. Thus we have

$$\begin{aligned}
(T\varrho)(v) &= \sum_{t(e)=v} \alpha^{h(\pi(o(e)))}\nu^{-\operatorname{ord}_{o(e)}(e_1)} \\
&= \alpha^{h(\pi(v))+1}\nu^{-\operatorname{ord}_v e_1} + q\cdot\alpha^{h(\pi(v))-1}\nu^{1-\operatorname{ord}_v(e_1)} \\
&= (\alpha + q\alpha^{-1}\nu)\alpha^{h(\pi(v))}\nu^{-\operatorname{ord}_v e_1} = a\varrho(v),
\end{aligned}$$

and also $(T\varrho)(\varpi^i v) = (TN^{-i}\varrho)(v) = N^{-i}(a\varrho)(v) = a\varrho(\varpi^i v)$ for a general $\varpi^i v \in \tilde{\mathcal{V}}$, which shows that $\varrho \in \ker(T-a)$. □

If $a^2 \neq \nu(q+1)^2$ (the "spherical case"), we put $\mathcal{B}_{a,\nu}(F,R) := \tilde{\mathcal{B}}_{a,\nu}(F,R)$ and $\mathcal{B}^{a,\nu}(F,M) := \tilde{\mathcal{B}}^{a,\nu}(F,M)$.

In the "special case" $a^2 = \nu(q+1)^2$, we need to assume that the polynomial $X^2 - a\nu X + q\nu^{-1} \in R[X]$ has a zero $\alpha' \in R$. Then the map $\varrho := \varrho_{\alpha',\nu} \in C(\tilde{\mathcal{V}},R)$ defined as above lies in $\tilde{\mathcal{B}}^{a\nu,\nu^{-1}}(F,R) = \ker(TN-a)\cap\ker(N^{-1}-\nu)$ by Lemma 2.7, since $a\nu = \alpha' + q\nu^{-1}/\alpha'$. In other words, the kernel of the map

$\langle \cdot, \varrho\rangle : C_c(\tilde{\mathcal{V}}, R) \to R$ contains $\operatorname{Im}(T-a) + \operatorname{Im}(N-\nu)$; and we define

$$\mathcal{B}_{a,\nu}(F,R) := \ker\left(\langle \cdot, \varrho\rangle\right) / \left(\operatorname{Im}(T-a) + \operatorname{Im}(N-\nu)\right)$$

to be the quotient; evidently, it is an R-submodule of codimension 1 of $\tilde{\mathcal{B}}_{a,\nu}(F,R)$. Dually, $T-a$ and $N-\nu$ both map the submodule $\varrho M = \{\varrho \cdot m, m \in M\}$ of $C(\tilde{\mathcal{V}}, M)$ to zero and thus induce endomorphisms on $C(\tilde{\mathcal{V}}, M)/\varrho M$; we define $\mathcal{B}^{a,\nu}(F,M)$ to be the intersection of their kernels.
In the special case, since $\nu = \alpha^2$, Lemma 2.7 states that $\varrho(gv_0) = \chi_\alpha(ad)\varrho(v_0) = \chi_\alpha(\det g)\varrho(v_0)$ for all $g = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in B(F)$, and thus for all $g \in G(F)$ by the Iwasawa decomposition, since $G(\mathcal{O}_F)$ fixes v_0 and lies in the kernel of $\chi_\alpha \circ \det$. By the multiplicity of det, we have $(g^{-1}\varrho)(v) = \varrho(gv) = \chi_\alpha(\det g)\varrho(v)$ for all $g \in G(F)$, $v \in \tilde{\mathcal{V}}$. So $\phi \in \ker\langle \cdot, \varrho\rangle$ implies $\langle g\phi, \varrho\rangle = \langle \phi, g^{-1}\varrho\rangle = \chi_\alpha(\det g)\langle \phi, \varrho\rangle = 0$, i.e. $\ker\langle \cdot, \varrho\rangle$ and thus $\mathcal{B}_{a,\nu}(F,R)$ are $G(F)$-modules.
By the adjointness properties of the Hecke operators T and N, we have pairings $\operatorname{coker}(T_c - a) \times \ker(TN - a) \to M$ and $\operatorname{coker}(N_c - \nu) \times \ker(N^{-1} - \nu) \to M$, which "combine" to give a pairing

$$\langle \text{-}, \text{-} \rangle : \mathcal{B}_{a,\nu}(F,R) \times \mathcal{B}^{a\nu,\nu^{-1}}(F,M) \to M$$

(since $\ker(TN-a) \cap \ker(N^{-1}-\nu) = \ker(T - a\nu) \cap \ker(N - \nu^{-1})$), and a corresponding isomorphism $\mathcal{B}^{a\nu,\nu^{-1}}(F,M) \xrightarrow{\cong} \operatorname{Hom}(\mathcal{B}_{a,\nu}(F,R), M)$.

Definition 2.8. Let G be a totally disconnected locally compact group, $H \subseteq G$ an open subgroup. For a smooth $R[H]$-module M, we define the *(compactly) induced G-representation* of M, denoted $\operatorname{Ind}_H^G M$, to be the space of maps $f : G \to M$ such that $f(hg) = f(g)$ for all $g \in G, h \in H$, and such that f has compact support modulo H. We let G act on $\operatorname{Ind}_H^G M$ via $g \cdot f(x) := f(xg)$. (We can also write $\operatorname{Ind}_H^G M = R[G] \otimes_{R[H]} M$, cf. [Br82], III.5.)
We further define $\operatorname{Coind}_H^G M := \operatorname{Hom}_{R[H]}(R[G], M)$. Finally, for an $R[G]$-module N, we write $\operatorname{res}_H^G N$ for its underlying $R[H]$-module ("restriction").

By Theorem 2.6, $T_c - a$ (as well as $N_c - \nu$) is injective, and the induced map

$$N_c - \nu : \operatorname{coker}(T_c - a) = C_c(\tilde{\mathcal{V}}, R)/\operatorname{Im}(T_c - a) \to \operatorname{coker}(T_c - a)$$

(of $R[T, N^{\pm 1}]/(T-a) = R[N^{\pm 1}]$-modules) is also injective. Now since $G(F)$ acts transitively on $\tilde{\mathcal{V}}$, with the stabilizer of $v_0 := \mathcal{O}_F^2$ being $K := G(\mathcal{O}_F)$, we have an isomorphism $C_c(\tilde{\mathcal{V}}, R) \cong \operatorname{Ind}_K^{G(F)} R$. Thus we have exact sequences

$$0 \to \operatorname{Ind}_K^{G(F)} R \xrightarrow{T-a} \operatorname{Ind}_K^{G(F)} R \to \operatorname{coker}(T_c - a) \to 0 \tag{4}$$

and (for a, ν in the spherical case)

$$0 \to \operatorname{coker}(T_c - a) \xrightarrow{N-\nu} \operatorname{coker}(T_c - a) \to \mathcal{B}_{a,\nu}(F,R) \to 0, \tag{5}$$

with all entries being free R-modules. Applying $\operatorname{Hom}_R(\cdot, M)$ to them, we get:

Lemma 2.9. *We have exact sequences of R-modules*

$$0 \to \ker(TN - a) \to \operatorname{Coind}_K^{G(F)} M \xrightarrow{T-a} \operatorname{Coind}_K^{G(F)} M \to 0$$

and, if $\mathcal{B}_{a,\nu}(F, M)$ is spherical (i.e. $a^2 \neq \nu(q+1)^2$),

$$0 \to \mathcal{B}^{a\nu,\nu^{-1}}(F, M) \to \ker(TN - a) \xrightarrow{N-\nu} \ker(TN - a) \to 0.$$

For the special case, we have to work a bit more to get similar exact sequences: By [Sp14], eq. (22), for the representation $St^-(F, R) := \mathcal{B}_{-(q+1),1}(F, R)$ (i.e. $\nu = 1$, $\alpha = -1$) with trivial central character, we have an exact sequence of G-modules

$$0 \to \operatorname{Ind}_{KZ}^G R \to \operatorname{Ind}_{K'Z}^G R \to St^-(F, R) \to 0, \tag{6}$$

where $K' = \langle W \rangle K_0(\mathfrak{p})$ is the subgroup of KZ generated by $W := \left(\begin{smallmatrix} 0 & 1 \\ \varpi & 0 \end{smallmatrix}\right)$ and the subgroup $K_0(\mathfrak{p}) \subseteq K$ of matrices that are upper-triangular modulo $\mathfrak{p}$. (Since $W^2 \in Z$, $K_0(\mathfrak{p})Z$ is a subgroup of K' of order 2.) Now aany special representation (π, V) can be written as $\pi = \chi \otimes St^-$ for some character $\chi = \chi_Z$ (cf. the proof of lemma 2.12 below), and is obviously G-isomorphic to the representation $\pi \otimes (\chi \circ \det)$ acting on the space $V \otimes_R R(\chi \circ \det)$, where $R(\chi \circ \det)$ is the ring R with G-module structure given via $gr = \chi(\det(g))r$ for $g \in G, r \in R$. Tensoring (6) with $R(\chi \circ \det)$ over R gives an exact sequence of G-modules

$$0 \to \operatorname{Ind}_{KZ}^G \chi \to \operatorname{Ind}_{K'Z}^G \chi \to V \to 0. \tag{7}$$

It is easily seen that $R(\chi \circ \det)$ fits into another exact sequence of G-modules

$$0 \to \operatorname{Ind}_H^G R \xrightarrow{\left(\begin{smallmatrix} \varpi & 0 \\ 0 & 1 \end{smallmatrix}\right) - \chi(\varpi)\,\mathrm{id}} \operatorname{Ind}_H^G R \xrightarrow{\psi} R(\chi \circ \det) \to 0,$$

where $H := \{g \in G | \det g \in \mathcal{O}_F^\times\}$ is a normal subgroup containing K, $\left(\begin{smallmatrix} \varpi & 0 \\ 0 & 1 \end{smallmatrix}\right)(f)(g) := f(\left(\begin{smallmatrix} \varpi & 0 \\ 0 & 1 \end{smallmatrix}\right)^{-1} g)$ for $f \in \operatorname{Ind}_H^G R = \{f : G \to R | f(Hg) = f(g)$ for all $g \in G\}$, $g \in G$, is the natural operation of G, and where ψ is the G-equivariant map defined by $1_U \mapsto 1$.

Now since $H \subseteq G$ is a normal subgroup, we have $\operatorname{Ind}_H^G R \cong R[G/H]$ as G-modules (in fact $G/H \cong \mathbb{Z}$ as an abstract group). Let $X \subseteq G$ be a subgroup such that the natural inclusion $X/(X \cap H) \hookrightarrow G/H$ has finite cokernel; let $g_i H$, $i = 1, \ldots n$ be a set of representatives of that cokernel. Then we have a (non-canonical) X-isomorphism $\bigoplus_{i=0}^n \operatorname{Ind}_{X \cap H}^X \to \operatorname{Ind}_H^G R$ defined via $(1_{(X \cap H)x})_i \mapsto 1_{Hxg_i}$ for each $i = 1, \ldots, n$ (cf. [Br82], III (5.4)).

Using this isomorphism and the "tensor identity" $\operatorname{Ind}_H^G M \otimes N \cong \operatorname{Ind}_H^G(M \otimes \operatorname{res}_H^G N)$ for any groups $H \subseteq G$, H-module M and G-module N ([Br82] III.5, Ex. 2), we have

$$\begin{aligned} \operatorname{Ind}_{KZ}^G R \otimes_R \operatorname{Ind}_H^G R &\cong \operatorname{Ind}_{KZ}^G(\operatorname{res}_{KZ}^G(\operatorname{Ind}_H^G R)) \\ &= \operatorname{Ind}_{KZ}^G((\operatorname{Ind}_{KZ \cap H}^{KZ} R)^2) \\ &= (\operatorname{Ind}_{KZ}^G(\operatorname{Ind}_K^{KZ} R))^2 = (\operatorname{Ind}_K^G R)^2 \end{aligned}$$

(since $KZ/KZ \cap H \hookrightarrow G/H$ has index 2), and similarly

$$\mathrm{Ind}_{K'Z}^{G} R \otimes_R \mathrm{Ind}_{H}^{G} R \cong (\mathrm{Ind}_{K'}^{G} R)^2.$$

Thus, we can resolve the first and second term of (7) into exact sequences

$$0 \to (\mathrm{Ind}_{K}^{G} R)^2 \to (\mathrm{Ind}_{K}^{G} R)^2 \to \mathrm{Ind}_{KZ}^{G} \chi \to 0,$$

$$0 \to (\mathrm{Ind}_{K'}^{G} R)^2 \to (\mathrm{Ind}_{K'}^{G} R)^2 \to \mathrm{Ind}_{\langle W \rangle K_0(\mathfrak{p}) Z}^{G} \chi \to 0.$$

Dualizing (7) and these by taking $\mathrm{Hom}(\cdot, M)$ for an R-module M, we get a "resolution" of $\mathcal{B}^{a\nu,\nu^{-1}}(F, M)$ in terms of coinduced modules:

Lemma 2.10. *We have exact sequences*

$$\begin{aligned}
&0 \to \mathcal{B}^{a\nu,\nu^{-1}}(F, M) \to \mathrm{Coind}_{K'Z}^{G} M(\chi) \to \mathrm{Coind}_{KZ}^{G} M(\chi) \to 0,\\
&0 \to \mathrm{Coind}_{KZ}^{G} M(\chi) \to (\mathrm{Coind}_{K}^{G} R)^2 \to (\mathrm{Coind}_{K}^{G} R)^2 \to 0,\\
&0 \to \mathrm{Coind}_{K'Z}^{G} M(\chi) \to (\mathrm{Coind}_{K'}^{G} R)^2 \to (\mathrm{Coind}_{K'}^{G} R)^2 \to 0
\end{aligned}$$

for all special $\mathcal{B}_{a,\nu}(F, R)$ (i.e. $a^2 = \nu(q+1)^2$), where $\chi = \chi_Z$ is the central character.

It is easily seen that the above arguments could be modified to get a similar set of exact sequences in the spherical case as well (replacing K' by K everywhere), in addition to that given in lemma 2.9; but we will not need this.

2.4 Distributions on the Bruhat-Tits graph

For $\varrho \in C(\tilde{\mathcal{V}}, R)$ we define R-linear maps

$$\tilde{\delta}_\varrho : C(\tilde{\mathcal{E}}, M) \to C(\tilde{\mathcal{V}}, M), \quad \tilde{\delta}_\varrho(c)(v) := \sum_{v=t(e)} \varrho(o(e))c(e) - \sum_{v=o(e)} \varrho(t(e))c(e),$$

$$\tilde{\delta}^\varrho : C(\tilde{\mathcal{V}}, M) \to C(\tilde{\mathcal{E}}, M), \quad \tilde{\delta}^\varrho(\phi)(e) := \varrho(o(e))\phi(t(e)) - \varrho(t(e))\phi(o(e)).$$

One easily checks that these are adjoint with respect to the pairings (2) and (3), i.e. we have $\langle \tilde{\delta}_\varrho(c), \phi \rangle = \langle c, \tilde{\delta}^\varrho(\phi) \rangle$ for all $c \in C_c(\tilde{\mathcal{E}}, R)$, $\phi \in C(\tilde{\mathcal{V}}, M)$. We denote the maps corresponding to $\varrho \equiv 1$ by $\delta := \tilde{\delta}_1$, $\delta^* := \tilde{\delta}^1$.
For each ϱ, the map $\tilde{\delta}_\varrho$ fits into an exact sequence

$$C_c(\tilde{\mathcal{E}}, R) \xrightarrow{\tilde{\delta}_\varrho} C_c(\tilde{\mathcal{V}}, R) \xrightarrow{\langle \cdot, \varrho \rangle} R \to 0$$

but it is not injective in general: e.g. for $\varrho \equiv 1$, the map $\tilde{\mathcal{E}} \to R$ symbolized by

$$\begin{tikzcd} \cdot \arrow[r, "-1"] \arrow[d, "1"] & \cdot \arrow[d, "-1"] \\ \cdot \arrow[r, "1"] & \cdot \end{tikzcd}$$

(and zero outside the square) lies in $\ker \delta$.
The restriction $\delta^*|_{C_c(\tilde{\mathcal{V}},R)}$ to compactly supported maps is injective since $\tilde{\mathcal{T}}$ has no directed circles, and we have a surjective map

$$\operatorname{coker}\big(\delta^* : C_c(\tilde{\mathcal{V}},R) \to C_c(\tilde{\mathcal{E}},R)\big) \to C^0(\mathbb{P}^1(F),R)/R, \quad c \mapsto \sum_{e\in\tilde{\mathcal{E}}} c(e) 1_{U(\pi(e))}$$

(which corresponds to an isomorphism of the similar map on the Bruhat-Tits tree $\mathcal{T}$). Its kernel is generated by the functions $1_{\{e\}} - 1_{\{e'\}}$ for $e, e' \in \tilde{\mathcal{E}}$ with $\pi(e) = \pi(e')$.
For $\varrho_1, \varrho_2 \in C(\tilde{\mathcal{V}}, R)$ and $\phi \in C(\tilde{\mathcal{V}}, M)$ it is easily checked that

$$\big(\tilde{\delta}_{\varrho_1} \circ \tilde{\delta}^{\varrho_2}\big)(\phi) = (T+TN)(\varrho_1 \cdot \varrho_2)\cdot\phi - \varrho_2 \cdot (T+TN)(\varrho_1\cdot\phi).$$

For $a' \in R$ and $\varrho \in \ker(T+TN-a')$, applying this equality for $\varrho_1 = \varrho$ and $\varrho_2 = 1$ shows that $\tilde{\delta}_\varrho$ maps $\operatorname{Im}\delta^*$ into $\operatorname{Im}(T+TN-a')$, so we get an R-linear map

$$\tilde{\delta}_\varrho : \operatorname{coker}\big(\delta^* : C_c(\tilde{\mathcal{V}},R) \to C_c(\tilde{\mathcal{E}},R)\big) \to \operatorname{coker}(T_c + T_cN_c - a').$$

Let now again $\alpha, \nu \in R^*$, and $a := \alpha + q\nu/\alpha$. We let $\varrho := \varrho_{\alpha,\nu} \in \tilde{\mathcal{B}}^{a,\nu}(F,R)$ as defined in lemma 2.7, and write $\tilde{\delta}_{\alpha,\nu} := \tilde{\delta}_\varrho$. Since $\tilde{\delta}_{\alpha,\nu}$ maps $1_{\{e\}} - 1_{\{\varpi e\}}$ into $\operatorname{Im}(R-\nu)$, it induces a map

$$\tilde{\delta}_{\alpha,\nu} : C^0(\mathbb{P}^1(F),R)/R \to \mathcal{B}_{a,\nu}(F,R)$$

(same name by abuse of notation) via the commutative diagram

$$\begin{array}{ccc}
\operatorname{coker}\delta^* & \xrightarrow{\tilde{\delta}_{\alpha,\nu}} & \operatorname{coker}(T_c + T_cN_c - a') \\
\downarrow & & \downarrow{\scriptstyle \bmod (N-\nu)} \\
C^0(\mathbb{P}^1(F),R)/R & \xrightarrow{\tilde{\delta}_{\alpha,\nu}} & \mathcal{B}_{a,\nu}(F,R)
\end{array}$$

with $a' := a(1+\nu)$, since $\varrho \in \mathcal{B}^{a,\nu}(F,R) \subseteq \ker(T+TN-a')$.

LEMMA 2.11. *We have $\varrho(gv) = \chi_\alpha(d/a')\chi_\nu(a')\varrho(v)$, and thus*

$$\tilde{\delta}_{\alpha,\nu}(gf) = \chi_\alpha(d/a')\chi_\nu(a') g \tilde{\delta}_{\alpha,\nu}(f),$$

for all $v \in \tilde{\mathcal{V}}$, $f \in C^0(\mathbb{P}^1(F),R)/R$ and $g = \begin{pmatrix} a' & b \\ 0 & d \end{pmatrix} \in B(F)$.

Proof. (a) Using lemma 2.5(b) and the fact that $\operatorname{ord}_{gv}(e_1) = -\operatorname{ord}_\varpi(a') + \operatorname{ord}_v(e_1)$, we have

$$\varrho\left(\begin{pmatrix} a' & b \\ 0 & d \end{pmatrix} v\right) = \alpha^{h(v) - \operatorname{ord}_\varpi(a'/d)} \nu^{\operatorname{ord}_\varpi(a') - \operatorname{ord}_v(e_1)} = \chi_\alpha(d/a')\chi_\nu(a')\varrho(v)$$

for all $v \in \tilde{\mathcal{V}}$. For f and g as in the assertion, we thus have

$$\begin{aligned}
\tilde{\delta}_{\alpha,\nu}(gf)(v) &= \sum_{v=t(e)} \varrho(o(e))f(g^{-1}e) - \sum_{v=o(e)} \varrho(t(e))f(g^{-1}e) \\
&= \sum_{g^{-1}v=t(e)} \varrho(o(ge))f(e) - \sum_{g^{-1}v=o(e)} \varrho(t(ge))f(e) \\
&= \chi_\alpha(d/a')\chi_\nu(a')\varrho(v)\Big(\sum_{g^{-1}v=t(e)} \varrho(o(e))f(e) - \sum_{g^{-1}v=o(e)} \varrho(t(e))f(e)\Big) \\
&= \chi_\alpha(d/a')\chi_\nu(a')g\tilde{\delta}_{\alpha,\nu}(f)(v).
\end{aligned}$$

□

We define a function $\delta_{\alpha,\nu} : C_c(F^*, R) \to \mathcal{B}_{a,\nu}(F, R)$ as follows: For $f \in C_c(F^*, R)$, we let $\psi_0(f) \in C_c(\mathbb{P}^1(F), R)$ be the extension of $x \mapsto \chi_\alpha(x)\chi_\nu(x)^{-1}f(x)$ by zero to $\mathbb{P}^1(F)$. We set $\delta_{\alpha,\nu} := \tilde{\delta}_{\alpha,\nu} \circ \psi_0$. If $\alpha = \nu$, we can define $\delta_{\alpha,\nu}$ on all functions in $C_c(F, R)$.
We let F^* operate on $C_c(F, R)$ by $(tf)(x) := f(t^{-1}x)$; this induces an action of the group $T^1(F) := \{\left(\begin{smallmatrix} t & 0 \\ 0 & 1 \end{smallmatrix}\right) | t \in F^*\}$, which we identify with F^* in the obvious way. With respect to it, we have

$$\psi_0(tf)(x) = \chi_\alpha(t)\chi_\nu(t)^{-1}t\psi_0(f)(x)$$

and

$$\tilde{\delta}_{\alpha,\nu}(tf) = \chi_\alpha^{-1}(t)\chi_\nu(t)t\tilde{\delta}_{\alpha,\nu}(f),$$

so $\delta_{\alpha,\nu}$ is $T^1(F)$-equivariant.
For an R-module M, we define an F^*-action on $\mathrm{Dist}(F^*, M)$ by $\int f d(t\mu) := t\int(t^{-1}f)d\mu$. Let $H \subseteq G(F)$ be a subgroup, and M an $R[H]$-module. We define an H-action on $\mathcal{B}^{a\nu,\nu^{-1}}(F, M)$ by requiring $\langle\phi, h\lambda\rangle = h \cdot \langle h^{-1}\phi, \lambda\rangle$ for all $\phi \in \mathcal{B}_{a,\nu}(F, M)$, $\lambda \in \mathcal{B}^{a\nu,\nu^{-1}}(F, M)$, $h \in H$. With respect to these two actions, we get a $T^1(F) \cap H$-equivariant mapping

$$\delta^{\alpha,\nu} : \mathcal{B}^{a\nu,\nu^{-1}}(F, M) \to \mathrm{Dist}(F^*, M), \quad \delta^{\alpha,\nu}(\lambda) := \langle\delta_{\alpha,\nu}(\cdot), \lambda\rangle$$

dual to $\delta_{\alpha,\nu}$.

2.5 Local distributions

Now consider the case $R = \mathbb{C}$. Let $\chi_1, \chi_2 : F^* \to \mathbb{C}^*$ be two unramified characters. We consider (χ_1, χ_2) as a character on the torus $T(F)$ of $\mathrm{GL}_2(F)$, which induces a character χ on $B(F)$ by

$$\chi\begin{pmatrix} t_1 & u \\ 0 & t_2 \end{pmatrix} := \chi_1(t_1)\chi_2(t_2).$$

Put $\alpha_i := \chi_i(\varpi)\sqrt{q} \in \mathbb{C}^*$ for $i = 1, 2$. Set $\nu := \chi_1(\varpi)\chi_2(\varpi) = \alpha_1\alpha_2q^{-1} \in \mathbb{C}^*$, and $a := \alpha_1 + \alpha_2 = \alpha_i + q\nu/\alpha_i$ for either i. When a and ν are given by the α_i

like this, we will often write $\mathcal{B}_{\alpha_1,\alpha_2}(F,R) := \mathcal{B}_{a,\nu}(F,R)$ and $\mathcal{B}^{\alpha_1,\alpha_2}(F,M) := \mathcal{B}^{a\nu,\nu^{-1}}(F,M)$ (!) for its dual. In the special case $a^2 = \nu(q+1)^2$, we assume the χ_i to be sorted such that $\chi_1 = |\cdot|\chi_2$.

Let $\mathcal{B}(\chi_1,\chi_2)$ denote the space of continuous maps $\phi : G(F) \to \mathbb{C}$ such that

$$\phi\left(\begin{pmatrix} t_1 & u \\ 0 & t_2 \end{pmatrix} g\right) = \chi_{\alpha_1}(t_1)\chi_{\alpha_2}(t_2)|t_1|\phi(g) \tag{8}$$

for all $t_1, t_2 \in F^*$, $u \in F$, $g \in G(F)$. $G(F)$ operates canonically on $\mathcal{B}(\chi_1,\chi_2)$ by right translation (cf. [Bu98], Ch. 4.5). If $\chi_1\chi_2^{-1} \neq |\cdot|^{\pm 1}$, $\mathcal{B}(\chi_1,\chi_2)$ is a model of the spherical representation $\pi(\chi_1,\chi_2)$; if $\chi_1\chi_2^{-1} = |\cdot|^{\pm 1}$, the special representation $\pi(\chi_1,\chi_2)$ can be given as an irreducible subquotient of codimension 1 of $\mathcal{B}(\chi_1,\chi_2)$.[2]

LEMMA 2.12. *We have a G-equivariant isomorphism $\tilde{\mathcal{B}}_{a,\nu}(F,\mathbb{C}) \cong \mathcal{B}(\chi_1,\chi_2)$. It induces an isomorphism $\mathcal{B}_{a,\nu}(F,\mathbb{C}) \cong \pi(\chi_1,\chi_2)$ both for spherical and special representations.*

Proof. We choose a "central" unramified character $\chi_Z : F^* \to \mathbb{C}$ satisfying $\chi_Z^2(\varpi) = \nu$; then we have $\chi_1 = \chi_Z{\chi_0}^{-1}, \chi_2 = \chi_Z\chi_0$ for some unramified character χ_0. We set $a' := \sqrt{q}\left(\chi_0(\varpi)^{-1} + \chi_0(\varpi)\right)$, which satisfies $a = \chi_Z(\varpi)a'$.

For a representation (π, V) of $G(F)$, by [Bu98], Ex. 4.5.9, we can define another representation $\chi_Z \otimes \pi$ on V via

$$(g,v) \mapsto \chi_Z(\det(g))\pi(g)v \qquad \text{for all } g \in G(F), v \in V,$$

and with this definition we have $\mathcal{B}(\chi_1,\chi_2) \cong \chi_Z \otimes \mathcal{B}(\chi_0^{-1},\chi_0)$. Since $\mathcal{B}(\chi_0^{-1},\chi_0)$ has trivial central character, [BL95], Thm. 20 (as quoted in [Sp14]) states that $\mathcal{B}(\chi_0^{-1},\chi_0) \cong \mathcal{B}_{a',1}(F,\mathbb{C}) \cong \operatorname{Ind}_{KZ}^{G(F)} R/\operatorname{Im}(T - a')$.

Define a G-linear map $\phi : \operatorname{Ind}_K^G R \to \chi_Z \otimes \operatorname{Ind}_{KZ}^G R$ by $1_K \mapsto (\chi_Z \circ \det) \cdot 1_{KZ}$. Since 1_K (resp. $(\chi_Z \circ \det) \cdot 1_{KZ}$) generates $\operatorname{Ind}_K^G R$ (resp. $\chi_Z \otimes \operatorname{Ind}_{KZ}^G R$) as a $\mathbb{C}[G]$-module, ϕ is well-defined and surjective.

ϕ maps $N1_K = \left(\begin{smallmatrix} \varpi & 0 \\ 0 & \varpi \end{smallmatrix}\right) 1_K$ to

$$\left(\begin{smallmatrix} \varpi & 0 \\ 0 & \varpi \end{smallmatrix}\right)((\chi_Z \circ \det) \cdot 1_{KZ}) = \chi_Z(\varpi)^2 \cdot ((\chi_Z \circ \det) \cdot 1_{KZ}) = \nu \cdot \phi(1_K).$$

Thus $\operatorname{Im}(N - \nu) \subseteq \ker\phi$, and in fact the two are equal, since the preimage of the space of functions of support in a coset KZg ($g \in G(F)$) under ϕ is exactly the space generated by the 1_{Kzg}, $z \in Z(F) = Z(\mathcal{O}_F)\{\left(\begin{smallmatrix} \varpi & 0 \\ 0 & \varpi \end{smallmatrix}\right)\}^{\mathbb{Z}}$.

Furthermore, ϕ maps $T1_K = \sum_{i \in \mathcal{O}_F/(\varpi) \cup \{\infty\}} N_i 1_K$ (with the N_i as in Lemma 2.7) to

$$\sum_i \chi_Z(\det(N_i)) \cdot ((\chi_Z \circ \det) \cdot N_i 1_{KZ}) = \chi_Z(\varpi) \cdot (\chi_Z \circ \det) T1_{KZ}$$

(since $\det(N_i) = \varpi$ for all i),and thus $\operatorname{Im}(T - a)$ is mapped to $\operatorname{Im}\left(\chi_Z(\varpi)T - a\right) = \operatorname{Im}\left(\chi_Z(\varpi)(T - a')\right) = \operatorname{Im}(T - a')$.

[2] Note that [Bu98] denotes this special representation by $\sigma(\chi_1,\chi_2)$, not by $\pi(\chi_1,\chi_2)$.

Putting everything together, we thus have G-isomorphisms

$$\begin{aligned} C_c(\tilde{\mathcal{V}},\mathbb{C})/\big(\mathrm{Im}(T-a)+\mathrm{Im}(N-\nu)\big) &\cong \mathrm{Ind}_K^G R/\big(\mathrm{Im}(T-a)+\mathrm{Im}(N-\nu)\big)\\ &\cong \chi_Z\otimes\big(\mathrm{Ind}_{KZ}^G R/\mathrm{Im}(T-a')\big)\\ &\cong \chi_Z\otimes\mathcal{B}(\chi_0^{-1},\chi_0)\cong\mathcal{B}(\chi_1,\chi_2). \end{aligned}$$

Thus, $\mathcal{B}_{a,\nu}(F,\mathbb{C})$ is isomorphic to the spherical principal series representation $\pi(\chi_1,\chi_2)$ for $a^2\neq\nu(q+1)^2$.
In the special case, $\mathcal{B}_{a,\nu}(F,\mathbb{C})$ is a G-invariant subspace of $\tilde{\mathcal{B}}_{a,\nu}(F,\mathbb{C})$ of codimension 1, so it must be mapped under the isomorphism to the unique G-invariant subspace of $\mathcal{B}(\chi_1,\chi_2)$ of codimension 1 (in fact, the unique infinite-dimensional irreducible G-invariant subspace, by [Bu98], Thm. 4.5.1), which is the special representation $\pi(\chi_1,\chi_2)$. □

By [Bu98], section 4.4, there exists thus for all pairs a,ν a *Whittaker functional* λ on $\mathcal{B}_{a,\nu}(F,\mathbb{C})$, i.e. a nontrivial linear map $\lambda:\mathcal{B}_{a,\nu}(F,\mathbb{C})\to\mathbb{C}$ such that $\lambda\left(\left(\begin{smallmatrix}1&x\\0&1\end{smallmatrix}\right)\phi\right)=\psi(x)\lambda(\phi)$. It is unique up to scalar multiples.
From it, we furthermore get a *Whittaker model* $\mathcal{W}_{a,\nu}$ of $\mathcal{B}_{a,\nu}(F,\mathbb{C})$:

$$\mathcal{W}_{a,\nu}:=\{W_\xi:GL_2(F)\to\mathbb{C}\,|\,\xi\in\mathcal{B}_{a,\nu}(F,\mathbb{C})\},$$

where $W_\xi(g):=\lambda(g\cdot\xi)$ for all $g\in GL_2(F)$. (see e.g. [Bu98], Ch. 3, eq. (5.6).)
Now write $\alpha:=\alpha_1$ for short. Recall the distribution $\mu_{\alpha,\nu}=\psi(x)\chi_{\alpha/\nu}(x)dx\in\mathrm{Dist}(F^*,\mathbb{C})$. For $\alpha=\nu$, it extends to a distribution on F. We have the following generalization of [Sp14], Prop. 3.10:

Proposition 2.13. *(a) There exists a unique Whittaker functional $\lambda=\lambda_{a,\nu}$ on $\mathcal{B}_{a,\nu}(F,\mathbb{C})$ such that $\delta^{\alpha,\nu}(\lambda)=\mu_{\alpha,\nu}$.*
(b) For every $f\in C_c(F^*,\mathbb{C})$, there exists $W=W_f\in\mathcal{W}_{a,\nu}$ such that

$$\int_{F^*}(af)(x)\mu_{\alpha,\nu}(dx)=W_f\begin{pmatrix}a&0\\0&1\end{pmatrix}.$$

If $\alpha=\nu$, then for every $f\in C_c(F,\mathbb{C})$, there exists $W_f\in\mathcal{W}_{a,\nu}$ such that

$$\int_{F}(af)(x)\mu_{\alpha,\nu}(dx)=W_f\begin{pmatrix}a&0\\0&1\end{pmatrix}.$$

(c) Let $H\subseteq U=\mathcal{O}_F^\times$ be an open subgroup, and write $W_H:=W_{1_H}$. For every $f\in C_c^0(F^*,\mathbb{C})^H$ we have

$$\int_{F^*}f(x)\mu_{\alpha,\nu}(dx)=[U:H]\int_{F^*}f(x)W_H\begin{pmatrix}x&0\\0&1\end{pmatrix}d^\times x.$$

Proof. (a) By [Sp14], we have a Whittaker functional of the Steinberg representation given by the composite

$$St(F,\mathbb{C}):=C^0(\mathbb{P}^1(F),\mathbb{C})/\mathbb{C}\xrightarrow{\cong}C_c(F,\mathbb{C})\xrightarrow{\Lambda}\mathbb{C},\tag{9}$$

where the first map is the F-equivariant isomorphism

$$C^0(\mathbb{P}^1(F),\mathbb{C})/\mathbb{C} \to C_c(F,\mathbb{C}), \quad \phi \mapsto f(x) := \phi(x) - \phi(\infty),$$

(with F acting on $C_c(F,\mathbb{C})$ by $(x\cdot f)(y) := f(y-x)$, and on $C^0(\mathbb{P}^1(F),\mathbb{C})/\mathbb{C}$ by $x\phi := \left(\begin{smallmatrix}1 & x\\ 0 & 1\end{smallmatrix}\right)\phi$), and the second is

$$\Lambda : C_c(F,\mathbb{C}) \to \mathbb{C}, \quad f \mapsto \int_F f(x)\psi(x)dx.$$

Let now $\lambda : \mathcal{B}_{a,\nu}(F,\mathbb{C}) \to \mathbb{C}$ be a Whittaker functional of $\mathcal{B}_{a,\nu}(F,\mathbb{C})$. By lemma 2.11, for $u = \left(\begin{smallmatrix}1 & x\\ 0 & 1\end{smallmatrix}\right) \in B(F)$,

$$(\lambda \circ \tilde{\delta}_{\alpha,\nu})(u\phi) = \lambda(u\tilde{\delta}_{\alpha,\nu}(\phi)) = \psi(x)\lambda(\tilde{\delta}_{\alpha,\nu}(\phi)),$$

so $\lambda \circ \tilde{\delta}_{\alpha,\nu}$ is a Whittaker functional if it is not zero.
To describe the image of $\tilde{\delta}_{\alpha,\nu}$, consider the commutative diagram

$$\begin{tikzcd}
C_c(\tilde{\mathcal{E}},R) \arrow[r, "\tilde{\delta}_{\alpha,\nu}"] \arrow[d, "(10)"] & C_c(\tilde{\mathcal{V}},R) \arrow[d, "\phi\mapsto\phi\cdot\varrho"] & & \\
C_c(\tilde{\mathcal{E}},R) \arrow[r, "\delta"] & C_c(\tilde{\mathcal{V}},R) \arrow[r, "\langle\cdot,1\rangle"] & R \arrow[r] & 0
\end{tikzcd}$$

where the vertical maps are defined by

$$C_c(\tilde{\mathcal{E}},R) \to C_c(\tilde{\mathcal{E}},R), \quad c \mapsto \big(e \mapsto c(e)\varrho(o(e))\varrho(t(e))\big) \tag{10}$$

resp. by mapping ϕ to $v \mapsto \phi(v)\varrho(v)$; both are obviously isomorphisms.
Since the lower row is exact, we have $\operatorname{Im}\delta = \ker\langle\cdot,1\rangle =: C_c^0(\tilde{\mathcal{V}},R)$ and thus $\operatorname{Im}\tilde{\delta}_{\alpha,\nu} = \varrho^{-1}\cdot C_c^0(\tilde{\mathcal{V}},R)$.
Since $\lambda \neq 0$ and $\mathcal{B}_{a,\nu}(F,\mathbb{C})$ is generated by (the equivalence classes of) the $1_{\{v\}}$, $v \in \tilde{\mathcal{V}}$, there exists a $v \in \tilde{\mathcal{V}}$ such that $\lambda(1_{\{v\}}) \neq 0$. Let ϕ be this $1_{\{v\}}$, and let $u = \left(\begin{smallmatrix}1 & x\\ 0 & 1\end{smallmatrix}\right) \in B(F)$ such that $x \notin \ker\psi$. Then

$$\varrho\cdot(u\phi - \phi) = \varrho\cdot(1_{\{u^{-1}v\}} - 1_{\{v\}}) = \varrho(v)(1_{\{u^{-1}v\}} - 1_{\{v\}}) \in C_c^0(\tilde{\mathcal{V}},R)$$

by lemma 2.11, so $0 \neq u\phi - \phi \in \operatorname{Im}\tilde{\delta}_{\alpha,\nu}$, but $\lambda(u\phi-\phi) = \psi(x)\lambda(\phi) - \lambda(\phi) \neq 0$.
So $\lambda \circ \tilde{\delta}_{\alpha,\nu} \neq 0$ is indeed a Whittaker functional. By replacing λ by a scalar multiple, we can assume $\lambda \circ \tilde{\delta}_{\alpha,\nu} = (9)$.
Considering λ as an element of $\mathcal{B}^{a\nu,\nu^{-1}}(F,\mathbb{C}) \cong \operatorname{Hom}(\mathcal{B}_{a,\nu}(F,\mathbb{C}),\mathbb{C})$, we have

$$\begin{aligned}
\delta^{\alpha,\nu}(\lambda)(f) &= \langle\delta_{\alpha,\nu}(f),\lambda\rangle \\
&= \Lambda(\chi_\alpha\chi_\nu^{-1}f) \\
&= \int_{F^*}\chi_\alpha(x)\chi_\nu^{-1}(x)f(x)\psi(x)dx \\
&= \mu_{\alpha,\nu}(f).
\end{aligned}$$

(b), (c) follow from (a) as in [Sp14]. □

2.6 Semi-local theory

We can generalize many of the previous constructions to the semi-local case, considering all primes $\mathfrak{p}|p$ at once.
So let $F_1, \ldots, F_m$ be finite extensions of $\mathbb{Q}_p$, and for each i, let q_i be the number of elements of the residue field of F_i. We put $\underline{F} := F_1 \times \cdots \times F_m$.
Let R again be a ring, and $a_i \in R, \nu_i \in R^*$ for each $i \in \{1, \ldots, m\}$. Put $\underline{a} := (a_1, \ldots, a_m)$, $\underline{\nu} := (\nu_1, \ldots, \nu_m)$. We define $\mathcal{B}_{\underline{a},\underline{\nu}}(\underline{F}, R)$ as the tensor product

$$\mathcal{B}_{\underline{a},\underline{\nu}}(\underline{F}, R) := \bigotimes_{i=1}^{m} \mathcal{B}_{a_i,\nu_i}(F_i, R).$$

For an R-module M, we define $\mathcal{B}^{\underline{a\nu},\underline{\nu}^{-1}}(\underline{F}, M) := \mathrm{Hom}_R(\mathcal{B}_{\underline{a},\underline{\nu}}(\underline{F}, R), M)$; let

$$\langle \cdot, \cdot \rangle : \mathcal{B}_{\underline{a},\underline{\nu}}(\underline{F}, R) \times \mathcal{B}^{\underline{a\nu},\underline{\nu}^{-1}}(\underline{F}, M) \to M \tag{11}$$

denote the evaluation pairing.
We have an obvious isomorphism

$$\bigotimes_{i=1}^{m} C_c^0(F_i^*, R) \to C_c^0(\underline{F}^*, R), \quad \bigotimes_i f_i \mapsto \left((x_i)_{i=1,\ldots,m} \mapsto \prod_{i=1}^{m} f_i(x_i) \right). \tag{12}$$

Now when we have $\alpha_{i,1}, \alpha_{i,2} \in R^*$ such that $a_i = \alpha_{i,1} + \alpha_{i,2}$ and $\nu_i = \alpha_{i,1}\alpha_{i,2}q_i^{-1}$, we can define the $T^1(\underline{F})$-equivariant map

$$\delta_{\underline{\alpha}_{1,2}} := \delta_{\underline{\alpha_1},\underline{\nu}} : C_c^0(\underline{F}, R) \to \mathcal{B}_{\underline{a},\underline{\nu}}(\underline{F}, R)$$

as the inverse of (12) composed with $\bigotimes_{i=1}^{m} \delta_{\alpha_{i,1},\nu_i}$.
Again, we will often write $\mathcal{B}_{\underline{\alpha_1},\underline{\alpha_2}}(F, R) := \mathcal{B}_{\underline{a\nu},\underline{\nu}^{-1}}(F, R)$ and $\mathcal{B}^{\underline{\alpha_1},\underline{\alpha_2}}(F, M) := \mathcal{B}^{\underline{a\nu},\underline{\nu}^{-1}}(F, M)$.
If $H \subseteq G(F)$ is a subgroup, and M an $R[H]$-module, we define an H-action on $\mathcal{B}^{\underline{a\nu},\underline{\nu}^{-1}}(F, M)$ by requiring $\langle \phi, h\lambda \rangle = h \cdot \langle h^{-1}\phi, \lambda \rangle$ for all $\phi \in \mathcal{B}_{\underline{a},\underline{\nu}}(F, M)$, $\lambda \in \mathcal{B}^{\underline{a\nu},\underline{\nu}^{-1}}(F, M)$, $h \in H$, and get a $T^1(\underline{F}) \cap H$-equivariant mapping

$$\delta^{\underline{\alpha_1},\underline{\alpha_2}} : \mathcal{B}^{\underline{a\nu},\underline{\nu}^{-1}}(F, M) \to \mathrm{Dist}(\underline{F}^*, M), \quad \delta^{\underline{\alpha_1},\underline{\alpha_2}}(\lambda) := \langle \delta_{\underline{\alpha_1},\underline{\alpha_2}}(\cdot), \lambda \rangle.$$

Finally, we have a homomorphism

$$\begin{aligned} \bigotimes_{i=1}^{m} \mathcal{B}^{a_i\nu_i,\nu_i^{-1}}(F_i, R) &\xrightarrow{\cong} \bigotimes_{i=1}^{m} \mathrm{Hom}_R(\mathcal{B}_{a_i\nu_i,\nu_i^{-1}}(F_i, R), R) \\ &\to \mathrm{Hom}(\mathcal{B}_{a_1,\nu_1}(F_1, R), \mathrm{Hom}(\mathcal{B}_{a_2,\nu_2}(F_2, R), \mathrm{Hom}(\ldots, R))\ldots) \\ &\xrightarrow{\cong} \mathcal{B}^{\underline{a\nu},\underline{\nu}^{-1}}(F, R). \end{aligned} \tag{13}$$

where the second map is given by $\otimes_i f_i \mapsto (x_1 \mapsto (x_2 \mapsto (\ldots \mapsto \prod_i f_i(x_i))\ldots)$, and the last map by iterating the adjunction formula of the tensor product.

3 Cohomology classes and global measures

3.1 Definitions

From now on, let F denote a number field, with ring of integers $\mathcal{O}_F$. For each finite prime v, let $U_v := \mathcal{O}_v^*$. Let $\mathbb{A} = \mathbb{A}_F$ denote the ring of adeles of F, and $\mathbb{I} = \mathbb{I}_F$ the group of ideles of F. For a finite subset S of the set of places of F, we denote by $\mathbb{A}^S := \{x \in \mathbb{A}_F \,|\, x_v = 0 \ \forall v \in S\}$ the S-adeles and by $\mathbb{I}^S$ the S-ideles, and put $F_S := \prod_{v \in S} F_v$, $U_S := \prod_{v \in S} U_v$, $U^S := \prod_{v \notin S} U_v$ (if S contains all infinite places of F), and similarly for other global groups.
For ℓ a prime number or ∞, we write S_ℓ for the set of places of F above ℓ, and abbreviate the above notations to $\mathbb{A}^\ell := \mathbb{A}^{S_\ell}$, $\mathbb{A}^{p,\infty} := \mathbb{A}^{S_p \cup S_\infty}$, and similarly write $\mathbb{I}^p$, $\mathbb{I}^\infty$, F_p, F_∞, U^∞, U_p, $U^{p,\infty}$, $\mathbb{I}_\infty$ etc.
Let F have r real embeddings and s pairs of complex embeddings. Set $d := r + s - 1$. Let $\{\sigma_0, \dots, \sigma_{r-1}, \sigma_r, \dots, \sigma_d\}$ be a set of representatives of these embeddings (i.e. for $i \geq r$, choose one from each pair of complex embeddings), and denote by $\infty_0, \dots, \infty_d$ the corresponding archimedian primes of F. We let $S_\infty^0 := \{\infty_1, \dots, \infty_d\} \subseteq S_\infty$.
For each place v, let dx_v denote the associated self-dual Haar measure on F_v, and $dx := \prod_v dx_v$ the associated Haar measure on $\mathbb{A}_F$. We define Haar measures $d^\times x_v$ on F_v^* by $d^\times x_v := c_v \frac{dx_v}{|x_v|_v}$, where $c_v = (1 - \frac{1}{q_v})^{-1}$ for v finite, $c_v = 1$ for $v|\infty$. For $v|\infty$ complex, we use the decomposition $\mathbb{C}^* = \mathbb{R}_+^* \times S^1$ (with $S^1 = \{x \in \mathbb{C}^*; |x| = 1\}$) to write $d^\times x_v = d^\times r_v \; d\vartheta_v$ for variables r_v, ϑ_v with $r_v \in \mathbb{R}_+^*$, $\vartheta_v \in S^1$.
Let $S_1 \subseteq S_p$ be a set of primes of F lying above p, $S_2 := S_p - S_1$. Let R be a topological Hausdorff ring.

Definition 3.1. We define the module of continuous functions

$$\mathcal{C}(S_1, R) := C(F_{S_1} \times F_{S_2}^* \times \mathbb{I}^{p,\infty}/U^{p,\infty}, R);$$

and let $\mathcal{C}_c(S_1, R)$ be the submodule of all compactly supported $f \in \mathcal{C}(S_1, R)$. We write $\mathcal{C}^0(S_1, R)$, $\mathcal{C}_c^0(S_1, R)$ for the submodules of locally constant maps (or of continuous maps where R is assumed to have the discrete topology).We further define

$$\mathcal{C}_c^\flat(S_1, R) := \mathcal{C}_c(\varnothing, R) + \mathcal{C}_c^\flat(S_1, R) \subseteq \mathcal{C}_c^\flat(S_1, R)$$

to be the module of continuous compactly supported maps that are "constant near $(0_\mathfrak{p}, x^\mathfrak{p})$" for each $\mathfrak{p} \in S_1$.

Definition 3.2. For an R-module M, let $\mathcal{D}_f(S_1, M)$ denote the R-module of maps

$$\phi : \mathfrak{Co}(F_{S_1} \times F_{S_2}^*) \times \mathbb{I}_F^{p,\infty} \to M$$

that are $U^{p,\infty}$-invariant and such that $\phi(\cdot, x^{p,\infty})$ is a distribution for each $x^{p,\infty} \in \mathbb{I}_F^{p,\infty}$.

Since $\mathbb{I}_F^{p,\infty}/U^{p,\infty}$ is a discrete topological group, $\mathcal{D}_f(S_1, M)$ naturally identifies with the space of M-valued distributions on $F_{S_1} \times F_{S_2}^* \times \mathbb{I}_F^{p,\infty}/U^{p,\infty}$. So there exists a canonical R-bilinear map

$$\mathcal{D}_f(S_1, M) \times \mathcal{C}_c^0(S_1, R) \to M, \quad (\phi, f) \mapsto \int f\, d\phi, \tag{14}$$

which is easily seen to induce an isomorphism $\mathcal{D}_f(S_1, M) \cong \mathrm{Hom}_R(\mathcal{C}_c^0(S_1, R), M)$.

For a subgroup $E \subseteq F^*$ and an $R[E]$-module M, we let E operate on $\mathcal{D}_f(S_1, M)$ and $\mathcal{C}_c^0(S_1, R)$ by $(a\phi)(U, x^{p,\infty}) := a\phi(a^{-1}U, a^{-1}x^{p,\infty})$ and $(af)(x^\infty) := f(a^{-1}x^\infty)$ for $a \in E$, $U \in \mathfrak{Co}(F_{S_1} \times F_{S_2}^*)$, $x^{\cdot} \in \mathbb{I}_F^{\cdot}$; thus we have $\int (af)\, d(a\phi) = a \int f\, d\phi$ for all a, f, ϕ.

When $M = V$ is a finite-dimensional vector space over a p-adic field, we write $\mathcal{D}_f^b(S_1, V)$ for the subset of $\phi \in \mathcal{D}_f(S_1, V)$ such that ϕ is even a measure on $F_{S_1} \times F_{S_2} \times \mathbb{I}_F^{p,\infty}/U^{p,\infty}$.

Definition 3.3. For a $\mathbb{C}$-vector space V, define $\mathcal{D}(S_1, V)$ to be the set of all maps $\phi : \mathfrak{Co}(F_{S_1} \times F_{S_2}^*) \times \mathbb{I}^p \to V$ such that:

(i) ϕ is invariant under $F^\times$ and $U^{p,\infty}$.

(ii) For $x^p \in \mathbb{I}^p$, $\phi(\cdot, x^p)$ is a distribution of $F_{S_1} \times F_{S_2}$.

(iii) For all $U \in \mathfrak{Co}(F_{S_1} \times F_{S_2}^*)$, the map $\phi_U : \mathbb{I} = F_p^\times \times \mathbb{I}^p \to V, (x_p, x^p) \mapsto \phi(x_p U, x^p)$ is smooth, and rapidly decreasing as $|x| \to \infty$ and $|x| \to 0$.

We will need a variant of this last set: Let $\mathcal{D}'(S_1, V)$ be the set of all maps $\phi \in \mathcal{D}(S_1, V)$ that are "$(S^1)^s$-invariant", i.e. such that for all complex primes ∞_j of F and all $\zeta \in S^1 = \{x \in \mathbb{C}^*; |x| = 1\}$, we have

$$\phi(U, x^{p,\infty_j}, \zeta x_{\infty_j}) = \phi(U, x^{p,\infty_j}, x_{\infty_j}) \text{ for all } x^p = (x^{p,\infty_j}, x_{\infty_j}) \in \mathbb{I}^p.$$

There is an obvious surjective map

$$\mathcal{D}(S_1, V) \to \mathcal{D}'(S_1, V), \quad \phi \mapsto \left((U, x) \mapsto \int_{(S^1)^s} \phi(U, x) d\vartheta_r \cdots d\vartheta_{r+s-1} \right)$$

given by integrating over $(S^1)^s \subseteq (\mathbb{C}^*)^s \hookrightarrow \mathbb{I}_\infty$.

Let F_+^* denote the set of all $x \in F*$ that are totally positive, i.e. positive with respect to every real embedding of F. (For F totally imaginary, we have $F^* = F_+^*$.) Let $F^{*\prime} \subseteq F_+^*$ be a maximal torsion-free subgroup of F_+^*. If F has at least one real embedding, we obviously have $F^{*\prime} = F_+^*$; for totally imaginary F, $F^{*\prime}$ is a subgroup of finite index of F^* with $F/F^{*\prime} \cong \mu_F$, the roots of unity of F.

We set

$$E' := F^{*\prime} \cap O_F^\times \subseteq O_F^\times,$$

so E' is a torsion-free $\mathbb{Z}$-module of rank d. E' operates freely and discretely on the space

$$\mathbb{R}_0^{d+1} := \left\{(x_0,\dots,x_d) \in \mathbb{R}^{d+1} | \sum_{i=0}^{d} x_i = 0\right\}$$

via the embedding

$$\begin{array}{rcl} E' & \hookrightarrow & \mathbb{R}_0^{d+1} \\ a & \mapsto & (\log|\sigma_i(a)|)_{i\in S_\infty} \end{array}$$

(cf. proof of Dirichlet's unit theorem, e.g. in [Neu92], Ch. 1), and the quotient $\mathbb{R}_0^{d+1}/E'$ is compact. We choose the orientation on $\mathbb{R}_0^{d+1}$ induced by the natural orientation on $\mathbb{R}^d$ via the isomorphism $\mathbb{R}^d \cong \mathbb{R}_0^{d+1}$, $(x_1,\dots,x_d) \mapsto (-\sum_{i=1}^d x_i, x_1,\dots,x_d)$. So $\mathbb{R}_0^{d+1}/E'$ becomes an oriented compact d-dimensional manifold.
Let $\mathcal{G}_p$ be the Galois group of the maximal abelian extension of F which is unramified outside p and ∞; for a $\mathbb{C}$-vector space V, let $\mathrm{Dist}(\mathcal{G}_p, V)$ be the set of V-valued distributions of $\mathcal{G}_p$. Denote by $\varrho : \mathbb{I}_F/F^* \to \mathcal{G}_p$ the projection given by global reciprocity.

3.2 Global measures

Now let $V = \mathbb{C}$, equipped with the trivial $F^{*\prime}$-action. We want to construct a commutative diagram

$$\begin{tikzcd} \mathcal{D}(S_1,\mathbb{C}) \arrow[rr, "\phi\mapsto\kappa_\phi"] \arrow[dr, "\phi\mapsto\mu_\phi"'] & & H^d\big(F^{*\prime}, \mathcal{D}_f(S_1,\mathbb{C})\big) \arrow[dl, "\kappa\mapsto\mu_\kappa=\kappa\cap\partial(\cdot)"] \\ & \mathrm{Dist}(\mathcal{G}_p,\mathbb{C}) & \end{tikzcd} \tag{15}$$

First, let R be any topological Hausdorff ring. Let $\overline{E'}$ denote the closure of E' in U_p. The projection map $\mathrm{pr} : \mathbb{I}^\infty/U^{p,\infty} \to \mathbb{I}^\infty/(\overline{E'} \times U^{p,\infty})$ induces an isomorphism

$$\mathrm{pr}^* : C_c(\mathbb{I}^\infty/(\overline{E'} \times U^{p,\infty}), R) \to H^0(E', C_c(\mathbb{I}^\infty/U^{p,\infty}, R)),$$

and the reciprocity map induces a surjective map $\overline{\varrho} : \mathbb{I}^\infty/(\overline{E'} \times U^{p,\infty}) \to \mathcal{G}_p$. Now we can define a map

$$\begin{aligned} \varrho^\sharp : H_0(F^{*\prime}/E', C_c(\mathbb{I}^\infty/(\overline{E'} \times U^{p,\infty}), R)) &\to C(\mathcal{G}_p, R), \\ [f] &\mapsto \Big(\overline{\varrho}(x) \mapsto \sum_{\zeta\in F^{*\prime}/E'} f(\zeta x) \text{ for } x \in \mathbb{I}^\infty/(\overline{E'} \times U^{p,\infty})\Big). \end{aligned}$$

This is an isomorphism, with inverse map $f \mapsto [(f\circ\overline{\varrho})\cdot 1_{\mathcal{F}}]$, where $1_{\mathcal{F}}$ is the characteristic function of a fundamental domain $\mathcal{F}$ of the action of $F^{*\prime}/E'$ on $\mathbb{I}^\infty/U^\infty$.

We get a composite map

$$\begin{aligned} C(\mathcal{G}_p, R) &\xrightarrow{(\varrho^\sharp)^{-1}} H_0\big(F^{*\prime}/E', C_c(\mathbb{I}^\infty/(\overline{E'} \times U^{p,\infty}), R)\big) \\ &\xrightarrow{\mathrm{pr}^*} H_0\big(F^{*\prime}/E', H^0(E', C_c(\mathbb{I}^\infty/U^{p,\infty}, R))\big) \\ &\longrightarrow H_0\big(F^{*\prime}/E', H^0(E', \mathcal{C}_c(S_1, R))\big), \end{aligned} \tag{16}$$

where the last arrow is induced by the "extension by zero" from $C_c(\mathbb{I}^\infty/U^{p,\infty}, R)$ to $\mathcal{C}_c(S_1, R)$.
Now let $\eta \in H_d(E', \mathbb{Z}) \cong \mathbb{Z}$ be the generator that corresponds to the given orientation of $\mathbb{R}_0^{d+1}$. This gives us, for every R-module A, a homomorphism

$$H_0\big(F^{*\prime}/E', H^0(E', A)\big) \xrightarrow{\cap\eta} H_0\big(F^{*\prime}/E', H_d(E', A)\big)$$

Composing this with the edge morphism

$$H_0\big(F^{*\prime}/E', H_d(E', A)\big) \to H_d(F^{*\prime}, A) \tag{17}$$

(and setting $A = \mathcal{C}_c(S_1, R)$) gives a map

$$H_0\big(F^{*\prime}/E', H^0(E', \mathcal{C}_c(S_1, R))\big) \to H_d\big(F^{*\prime}, \mathcal{C}_c(S_1, R)\big) \tag{18}$$

We define

$$\partial : C(\mathcal{G}_p, R) \to H_d\big(F^{*\prime}, \mathcal{C}_c(S_1, R)\big)$$

as the composition of (16) with this map.
Now, letting M be an R-module equipped with the trivial $F^{*\prime}$-action, the bilinear form (14)

$$\begin{array}{rcl} \mathcal{D}_f(S_1, M) \times \mathcal{C}_c(S_1, R) & \to & M \\ (\phi, f) & \mapsto & \displaystyle\int f \, d\phi \end{array}$$

induces a cap product

$$\cap : H^d\big(F^{*\prime}, \mathcal{D}_f(S_1, M)\big) \times H_d\big(F^{*\prime}, \mathcal{C}_c(S_1, R)\big) \to H_0(F^{*\prime}, M) = M. \tag{19}$$

Thus for each $\kappa \in H^d(F^{*\prime}, \mathcal{D}_f(S_1, M))$, we get a distribution μ_κ on $\mathcal{G}_p$ by defining

$$\int_{\mathcal{G}_p} f(\gamma)\, \mu_\kappa(d\gamma) := \kappa \cap \partial(f) \tag{20}$$

for all continuous maps $f : \mathcal{G}_p \to R$.
Now let $M = V$ be a finite-dimensional vector space over a p-adic field K, and let $\kappa \in H^d(F^{*\prime}, \mathcal{D}_f^b(S_1, V))$. We identify κ with its image in $H^d(F^{*\prime}, \mathcal{D}_f(S_1, V))$; then it is easily seen that μ_κ is also a measure, i.e. we have a map

$$H^d(F^{*\prime}, \mathcal{D}_f^b(S_1, V)) \to \mathrm{Dist}^b(\mathcal{G}_p, V), \quad \kappa \mapsto \mu_\kappa. \tag{21}$$

Let $L|F$ be a $\mathbb{Z}_p$-extension of F. Since it is unramified outside p, it gives rise to a continuous homomorphism $\mathcal{G}_p \to \mathrm{Gal}(L|F)$ via $\sigma \mapsto \sigma|_L$. Fixing an isomorphism $\mathrm{Gal}(L|F) \cong p^{\varepsilon_p}\mathbb{Z}_p$ (where $\varepsilon_p = 2$ for $p = 2$, $\varepsilon_p = 1$ for p odd), we obtain a surjective homomorphism $\ell : \mathcal{G}_p \to p^{\varepsilon_p}\mathbb{Z}_p$. (Note that $p^{\varepsilon_p}\mathbb{Z}_p$ is the space of definition of the p-adic exponential function $\exp_p$.)

Example 3.4. Let L be the cyclotomic $\mathbb{Z}_p$-extension of F. Then we can take $\ell = \log_p \circ \mathcal{N}$, where $\mathcal{N} : \mathcal{G}_p \to \mathbb{Z}_p^*$ is the p-adic cyclotomic character, defined by requiring $\gamma\zeta = \zeta^{\mathcal{N}(\gamma)}$ for all $\gamma \in \mathcal{G}_p$ and all p-power roots of unity ζ. It is well-known (cf. [Wa82], par. 5) that $\log_p(\mathbb{Z}_p^*) = p^{\varepsilon_p}\mathbb{Z}_p$.

It is well-known that F has t independent $\mathbb{Z}_p$-extensions, where $s+1 \le t \le [F : \mathbb{Q}]$; the Leopoldt conjecture implies $t = s+1$. μ_κ defines a t-variable p-adic L-function as follows:

Definition 3.5. Let K be a p-adic field, V a finite-dimensional K-vector space, $\kappa \in H^d(F^{*\prime}, \mathcal{D}_f^b(S_1, V))$. Let $\ell_1, \ldots, \ell_t : \mathcal{G}_p \to p^{\varepsilon_p}\mathbb{Z}_p$ be continuous homomorphisms.The *p-adic L-function* of κ is given by

$$L_p(\underline{s}, \kappa) := L_p(s_1, \ldots, s_t, \kappa) := \int_{\mathcal{G}_p} \left(\prod_{i=1}^{t} \exp_p(s_i \ell_i(\gamma)) \right) \mu_\kappa(d\gamma)$$

for all $s_1, \ldots, s_t \in \mathbb{Z}_p$.

Remark 3.6. Let $\Sigma := \{\pm 1\}^r$, where r is the number of real embeddings of F. The group isomorphism $\mathbb{Z}/2\mathbb{Z} \cong \{\pm 1\}, \varepsilon \mapsto (-1)^\varepsilon$, induces a pairing

$$\langle \cdot, \cdot \rangle : \Sigma \to \{\pm 1\}, \quad \langle ((-1)^{\varepsilon_i})_i, ((-1)^{\varepsilon_i'})_i \rangle := (-1)^{\sum_i \varepsilon_i \varepsilon_i'}.$$

For a field k of characteristic zero, a $k[\Sigma]$-module V and $\underline{\mu} = (\mu_0, \ldots, \mu_{r-1}) \in \Sigma$, we put $V_{\underline{\mu}} := \{v \in V \,|\, \langle \underline{\mu}, \underline{\nu} \rangle v = \underline{\nu} v \; \forall \underline{\nu} \in \Sigma\}$, so that we have $V = \bigoplus_{\underline{\mu} \in \Sigma} V_{\underline{\mu}}$. We write $v_{\underline{\mu}}$ for the projection of $v \in V$ to $V_{\underline{\mu}}$, and $v_+ := v_{(1,\ldots,1)}$.
For $r > 0$, we identify Σ with $F^*/F^{*\prime}$ via the isomorphism $\Sigma \cong \prod_{i=0}^{r-1} \mathbb{R}^*/\mathbb{R}_+^* \cong F^*/F^{*\prime} = F^*/F_+^*$. Then for each F^*-module M, Σ acts on $H^d(F^{*\prime}, \mathcal{D}_f(S_1, M))$ and on $H^d(F^{*\prime}, \mathcal{D}_f^b(S_1, M))$. For $r = 0$, we let the trivial group Σ act on these groups as well for ease of notation. The exact sequence $\Sigma \cong \prod_{i=0}^{r-1} \mathbb{R}^*/\mathbb{R}_+^* = \mathbb{I}_\infty/\mathbb{I}_\infty^0 \to \mathcal{G}_p \to \mathcal{G}_p^+ \to 0$ of class field theory (where $\mathbb{I}_\infty^0$ is the maximal connected subgroup of $\mathbb{I}_\infty$) yields an action of Σ on $\mathcal{G}_p$. We easily check that (21) is Σ-equivariant, and that the maps $\gamma \mapsto \exp_p(s\ell_i(\gamma))$ factor over $\mathcal{G}_p \to \mathcal{G}_p^+$ (since $\mathbb{Z}_p$-extensions are unramified at ∞). Therefore we have $L_p(\underline{s}, \kappa) = L_p(\underline{s}, \kappa_+)$.

For $\phi \in \mathcal{D}(S_1, V)$ and $f \in C^0(\mathbb{I}/F^*, \mathbb{C})$, let

$$\int_{\mathbb{I}/F^*} f(x)\phi(d^\times x_p, x^p)\, d^\times x^p := [U_p : U] \int_{\mathbb{I}/F^*} f(x)\phi_U(x)\, d^\times x,$$

where we choose an open set $U \subseteq U_p$ such that $f(x_p u, x^p) = f(x_p, x^p)$ for all $(x_p, x^p) \in \mathbb{I}$ and $u \in U$; such a U exists by lemma 3.7 below. Since this integral is additive in f, there exists a unique V-valued distribution μ_ϕ on $\mathcal{G}_p$ such that

$$\int_{\mathcal{G}_p} f\, d\mu_\phi = \int_{\mathbb{I}/F^*} f(\varrho(x))\phi(d^\times x_p, x^p)\, d^\times x^p \tag{22}$$

for all functions $f \in C^0(\mathcal{G}_p, V)$.

Lemma 3.7. *Let $F : \mathbb{I}/F^* \to X$ be a locally constant map to a set X. Then there exists an open subgroup $U \subseteq \mathbb{I}$ such that f factors over $\mathbb{I}/F^*U$.*

Proof. $\mathbb{I}_\infty = \prod_{v|\infty} F_v$ is connected, thus f factors over $\overline{f} : \mathbb{I}/F^*\mathbb{I}_\infty \to X$. Since $\mathbb{I}/F^*\mathbb{I}_\infty$ is profinite, $\overline{f}$ further factors over a subgroup $U' \subseteq \mathbb{I}^\infty$ of finite index, which is open. □

Let $U^0_\infty := \prod_{v\in S^0_\infty} \mathbb{R}^*_+$; the isomorphisms $U^0_\infty \cong \mathbb{R}^d$, $(r_v)_v \mapsto (\log r_v)_v$, and $\mathbb{R}^d \cong \mathbb{R}^{d+1}_0$ give it the structure of a d-dimensional oriented manifold (with the natural orientation). It has the d-form $d^\times r_1 \cdot \ldots \cdot d^\times r_d$, where (by slight abuse of notation) we choose $d^\times r_i$ on F_{∞_i} corresponding to the Haar measure $d^\times x_i$ resp. $d^\times r_i$ on $\mathbb{R}^*_+ \subseteq F^*_{\infty_i}$. E' operates on U^0_∞ via $a \mapsto (|\sigma_i(a)|)_{i\in S^0_\infty}$, so the isomorphism $U^0_\infty \cong \mathbb{R}^{d+1}_0$ is E'-equivariant.
For $\phi \in \mathcal{D}'(S_1, V)$, set

$$\begin{aligned}\int_0^\infty \phi\, d^\times r_0 \colon \mathfrak{Co}(F_{S_1} \times F^*_{S_2}) \times \mathbb{I}^{p,\infty_0} &\to \mathbb{C}\\ (U, x^{p,\infty_0}) &\mapsto \int_0^\infty \phi(U, r_0, x^{p,\infty_0})\, d^\times r_0,\end{aligned}$$

where we let $r_0 \in F_{\infty_0}$ run through the positive real line $\mathbb{R}^*_+$ in F_{∞_0}. Composing this with the projection $\mathcal{D}(S_1, V) \to \mathcal{D}'(S_1, V)$ gives us a map

$$\begin{aligned}\mathcal{D}(S_1, V) &\to H^0\big(F^{*\prime}, \mathcal{D}_f(S_1, C^\infty(U^0_\infty, V))\big),\\ \phi &\mapsto \int_{(S^1)^s} \left(\int_0^\infty \phi\, d^\times r_0\right) d\vartheta_r\, d\vartheta_{r+1} \ldots d\vartheta_{r+s-1}\end{aligned} \tag{23}$$

(where $C^\infty(U^0_\infty, V)$ denotes the space of smooth V-valued functions on U^0_∞), since one easily checks that $\int_0^\infty \phi\, d^\times r_0$ is $F^{*\prime}$-invariant.
Define the complex $C^\bullet := \mathcal{D}_f(S_1, \Omega^\bullet(U^0_\infty, V))$. By the Poincare lemma, this is a resolution of $\mathcal{D}_f(S_1, V)$. We now define the map $\phi \mapsto \kappa_\phi$ as the composition of (23) with the composition

$$H^0\big(F^{*\prime}, \mathcal{D}_f(S_1, C^\infty(U^0_\infty, V))\big) \to H^0(F^{*\prime}, C^d) \to H^d(F^{*\prime}, \mathcal{D}_f(S_1, V)), \tag{24}$$

where the first map is induced by

$$C^\infty(U^0_\infty, V) \to \Omega^d(U^0_\infty, V), \quad f \mapsto f(r_1, \ldots, r_d) d^\times r_1 \cdot \ldots \cdot d^\times r_d, \tag{25}$$

and the second is an edge morphism in the spectral sequence

$$H^q(F^{*\prime}, C^p) \Rightarrow H^{p+q}(F^{*\prime}, \mathcal{D}_f(S_1, V)). \tag{26}$$

Specializing to $V = \mathbb{C}$, we now have:

PROPOSITION 3.8. *The diagram* (15) *commutes, i.e., for each* $\phi \in \mathcal{D}(S_1, \mathbb{C})$, *we have*

$$\mu_\phi = \mu_{\kappa_\phi}.$$

Proof. Analoguously to [Sp14], proof of prop. 4.21, we define a pairing

$$\langle\ ,\ \rangle \colon \mathcal{D}(S_1, \mathbb{C}) \times C^0(\mathcal{G}_p, \mathbb{C}) \to \mathbb{C}$$

as the composite of (23) $\times$ (16) with

$$\begin{aligned} H^0\big(F^{*\prime}, \mathcal{D}_f(S_1, C^\infty(U^0_\infty, \mathbb{C}))\big) \times H_0\big(F^{*\prime}/E', H^0(E', \mathcal{C}^0_c(S_1, \mathbb{C}))\big) \\ \xrightarrow{\cap} H_0\big(F^{*\prime}/E', H^0(E', C^\infty(U^0_\infty, \mathbb{C}))\big) \to H_0(F^{*\prime}/E', \mathbb{C}) \cong \mathbb{C}, \end{aligned} \tag{27}$$

where $\cap$ is the cap product induced by (14), and the second map is induced by

$$H^0\big(E', \mathcal{C}^\infty(U^0_\infty, \mathbb{C})\big) \to \mathbb{C}, \quad f \mapsto \int_{U^0_\infty/E'} f(r_1, \dots, r_d)\, d^\times r_1 \dots d^\times r_d. \tag{28}$$

Then we can show that

$$\kappa_\phi \cap \partial(f) = \langle \phi, f \rangle = \int_{\mathcal{G}_p} f(\gamma)\, \mu_\phi(d\gamma) \quad \text{for all } f \in C^0(\mathcal{G}_p, \mathbb{C}),$$

by copying the proof for the totally real case (replacing F^*_+ by $F^{*\prime}$, E_+ by E'), using the fact that for a d-form on the d-dimensional oriented manifold $M := \mathbb{R}^{d+1}_0/E' \cong U^0_\infty/E'$, integration over M corresponds to taking the cap product with the fundamental class η of M under the canonical isomorphism $H^d_{dR}(M) \cong H^d_{sing}(M) = H^d(E', \mathbb{C})$. ◻

3.3 EXCEPTIONAL ZEROS

Now let $\ell_1, \dots, \ell_t : \mathcal{G}_p \to \mathbb{Z}_p$ be continuous homomorphisms. Let again $S_1 = \{\mathfrak{p}_1, \dots, \mathfrak{p}_n\} \subseteq S_p$ be a set of primes above p, of cardinality $n := \#S_1$.

PROPOSITION 3.9. *For each* $\underline{x} = (x_1, \dots, x_t) \in \mathbb{N}_0^t$ *set* $|\underline{x}| := \sum_{i=1}^t x_i$. *Then*

$$\partial\Big(\prod_{i=1}^t \ell_i^{x_i}\Big) = 0 \quad \textit{for all } \underline{x} \textit{ with } |\underline{x}| \le n-1.$$

Proof. We can readily generalize the proof of Spieß' result for the p-adic cyclotomic character ($\ell = \log_p \circ \mathcal{N}$) in the totally real case ([Sp14], Prop. 4.6(a), Lemmas 4.1 and 4.7) to show that $\partial(\ell^x) = 0$ for all $0 \leq x \leq n-1$, using the facts that we can write $F^{*\prime} = E' \times \mathcal{T}$ for some subgroup $\mathcal{T} \subseteq F^{*\prime}$ (since $F^{*\prime}/E' = F^*/\mathcal{O}_F^\times$ is a free $\mathbb{Z}$-module), and that for each homomorphism $\ell : \mathcal{G}_p \to \mathbb{Z}_p$, the composition

$$\tilde{\ell} : \mathbb{I}^\infty \xrightarrow{\varrho} \mathcal{G}_p \xrightarrow{\ell} \mathbb{Z}_p \hookrightarrow \mathbb{Q}_p.$$

is zero on $\mathbb{I}^{\infty,p}$ (since the pro-q-part of $\mathcal{G}_p$ is finite for every prime $q \neq p$ and $\mathbb{Q}_p$ is torsion-free).

Now for a ring $R \supseteq \mathbb{Q}$, each monomial $\prod_{i=1}^t X_i^{n_i} \in R[X_1, \ldots, X_t]$ of degree $n = \sum_i n_i$ can be written as a linear combination of n-th powers $(X_i + r_{i,j}X_j)^n$. Therefore each product $\prod_{i=1}^t \ell_i^{x_i}$ of degree $x = |\underline{x}|$ is a linear combination of x-th powers of the homomorphisms $\ell_{i,j} := \ell_i + r_{i,j}\ell_j : \mathcal{G}_p \to \mathbb{Z}_p$. This proves the proposition. □

Definition 3.10. A t-variable p-adic analytic function $f(\underline{s}) = f(s_1, \ldots, s_t)$ ($s_i \in \mathbb{Z}_p$) *has vanishing order* $\geq n$ at the point $\underline{0} = (0, \ldots, 0)$ if all its partial derivatives of total order $\leq n-1$ vanish, i.e. if

$$\frac{\partial^k}{(\partial \underline{s})^{\underline{k}}} f(\underline{0}) := \frac{\partial^k}{\partial s_1^{k_1} \cdots \partial s_t^{k_t}} f(\underline{0}) = 0$$

for all $\underline{k} = (k_1, \ldots, k_t) \in \mathbb{N}_0^t$ with $k := |\underline{k}| \leq n-1$. We write $\operatorname{ord}_{\underline{s}=\underline{0}} f(\underline{s}) \geq n$ in this case.

Theorem 3.11. *Let $n := \#(S_1)$, $\kappa \in H^d(F^{*\prime}, \mathcal{D}_f^b(S_1, V))$, V a finite-dimensional vector space over a p-adic field. Then $L_p(\underline{s}, \kappa)$ is a locally analytic function, and we have*

$$\operatorname{ord}_{\underline{s}=\underline{0}} L_p(\underline{s}, \kappa) \geq n.$$

Proof. We have

$$\frac{\partial^k}{(\partial \underline{s})^{\underline{k}}} L_p(\underline{0}, \kappa) = \int_{\mathcal{G}_p} \left(\prod_{i=1}^t \ell_i(\gamma)^{k_i} \right) \mu_\kappa(d\gamma) = \kappa \cap \partial \left(\prod_{i=1}^t \ell_i(\gamma)^{k_i} \right)$$

for all $\underline{k} = (k_1, \ldots, k_t) \in \mathbb{N}_0^t$. Thus the theorem follows from proposition 3.9. □

3.4 Integral cohomology classes

Definition 3.12. A nonzero cohomology class $\kappa \in H^d(F^{*\prime}, \mathcal{D}_f(S_1, \mathbb{C}))$ is called *integral* if κ lies in the image of

$$H^d(F^{*\prime}, \mathcal{D}_f(S_1, R)) \otimes_R \mathbb{C} \to H^d(F^{*\prime}, \mathcal{D}_f(S_1, \mathbb{C}))$$

for some Dedekind ring $R \subseteq \overline{\mathcal{O}}$. If, in addition, there exists a torsion-free R-submodule $M \subseteq H^d(F^{*\prime}, \mathcal{D}_f(S_1, R))$ of rank ≤ 1 (i.e. M can be embedded into R) such that κ lies in the image of $M \otimes_R \mathbb{C} \to H^d(F^{*\prime}, \mathcal{D}_f(S_1, \mathbb{C}))$, then κ is *integral of rank* ≤ 1.

For κ as in def. 3.12 and $R \subseteq \mathbb{C}$, we let $L_{\kappa,R}$ be the image of

$$H_d(F^{*\prime}, \mathcal{C}_c^0(S_1, R)) \to H_0(F^{*\prime}, \mathbb{C}) = \mathbb{C}, \quad x \mapsto \kappa \cap x.$$

PROPOSITION 3.13. *Let $\kappa \in H^d(F^{*\prime}, \mathcal{D}_f(S_1, \mathbb{C}))$ be integral. Then*
(a) μ_κ is a p-adic measure.
(b) There exists a Dedekind ring $R \subseteq \overline{\mathcal{O}}$ such that $L_{\kappa,R}$ is a finitely generated R-module (resp. a torsion-free R-module of rank ≤ 1, if κ is integral of rank ≤ 1).
For each such R, the map $H^d(F^{*\prime}, \mathcal{D}_f(S_1, L_{\kappa,R})) \otimes \overline{\mathbb{Q}} \to \mathcal{H}^d(F^{*\prime}, \mathcal{D}_f(S_1, \mathbb{C}))$ is injective and κ lies in its image.

Proof. The proofs of the corresponding results for totally real F ([Sp14], prop. 4.17 and cor. 4.18) also work in the general case. □

Remark 3.14. Let κ be integral with Dedekind ring R as above. By (b) of the proposition, we can view κ as an element of $H^d(F^{*\prime}, \mathcal{D}_f(S_1, L_{\kappa,R})) \otimes \overline{\mathbb{Q}}$. Put $V_\kappa := L_{\kappa,R} \otimes_R \mathbb{C}_p$; let $\overline{\kappa}$ be the image of κ under the composition

$$\begin{aligned} H^d(F^{*\prime}, \mathcal{D}_f(S_1, L_{\kappa,R})) \otimes_R \overline{\mathbb{Q}} &\to H^d(F^{*\prime}, \mathcal{D}_f(S_1, L_{\kappa,R})) \otimes_R \mathbb{C}_p \\ &\to H^d(F^{*\prime}, \mathcal{D}_f^b(S_1, V_\kappa)), \end{aligned}$$

where the second map is induced by $\mathcal{D}_f(S_1, L_{\kappa,R}) \otimes_R \mathbb{C}_p \to \mathcal{D}_f^b(S_1, V_\kappa)$. By [Sp14], lemma 4.15, $\overline{\kappa}$ does not depend on the choice of R.
Since μ_κ is a p-adic measure, $\mu_{\overline{\kappa}}$ allows integration of all continuous functions $f \in C(\mathcal{G}_p, \mathbb{C}_p)$, and by abuse of notation, we write $L_p(\underline{s}, \kappa) := L_p(\underline{s}, \overline{\kappa})$ (cf. remark 3.6). So $L_p(\underline{s}, \kappa)$ has values in the finite-dimensional $\mathbb{C}_p$-vector space V_κ.

4 p-ADIC L-FUNCTIONS OF AUTOMORPHIC FORMS

We keep the notations from chapter 3; so F is again a number field with r real embeddings and s pairs of complex embeddings.
For an ideal $0 \neq \mathfrak{m} \subseteq \mathcal{O}_F$, we let $K_0(\mathfrak{m})_v \subseteq G(\mathcal{O}_{F_v})$ be the subgroup of matrices congruent to an upper triangular matrix modulo $\mathfrak{m}$, and we set $K_0(\mathfrak{m}) := \prod_{v \nmid \infty} K_0(\mathfrak{m})_v$, $K_0(\mathfrak{m})^S := \prod_{v \nmid \infty, v \notin S} K_0(\mathfrak{m})_v$ for a finite set of primes S. For each $\mathfrak{p}|p$, let $q_\mathfrak{p} = N(\mathfrak{p})$ denote the number of elements of the residue class field of $F_\mathfrak{p}$.
We denote by $|\cdot|_\mathbb{C}$ the square of the usual absolute value on $\mathbb{C}$, i.e. $|z|_\mathbb{C} = z\overline{z}$ for all $z \in \mathbb{C}$, and write $|\cdot|_\mathbb{R}$ for the usual absolute value on $\mathbb{R}$ in context. We write $|\alpha| := |\alpha|_\mathbb{C}^{\frac{1}{2}}$ for the archimedian absolute value when α is given as a complex number in the context; whereas in the context of the p-adic characters, $|\cdot|$ denotes the p-adic absolute value, unless otherwise noted.

Definition 4.1. Let $\mathfrak{A}_0(G, \underline{2}, \chi_Z)$ denote the set of all *cuspidal automorphic representations* $\pi = \otimes_v \pi_v$ of $G(\mathbb{A}_F)$ with central character χ_Z such that $\pi_v \cong \sigma(|\cdot|_{F_v}^{1/2}, |\cdot|_{F_v}^{-1/2})$ at all archimedian primes v. Here we follow the notation of [JL70]; so $\sigma(|\cdot|_{F_v}^{1/2}, |\cdot|_{F_v}^{-1/2})$ is the discrete series of weight 2, $\mathcal{D}(2)$, if v is real, and is isomorphic to the principal series representation $\pi(\mu_1, \mu_2)$ with $\mu_1(z) = z^{1/2}\overline{z}^{\,-1/2}$, $\mu_2(z) = z^{-1/2}\overline{z}^{1/2}$ if v is complex (cf. section 4.5 below).

We will only consider automorphic representations that are *p-ordinary* , i.e $\pi_\mathfrak{p}$ is ordinary (in the sense of chapter 2) for every $\mathfrak{p}|p$.
Therefore, for each $\mathfrak{p}|p$ we fix two non-zero elements $\alpha_{\mathfrak{p},1}, \alpha_{\mathfrak{p},2} \in \overline{\mathcal{O}} \subseteq \mathbb{C}$ such that $\pi_{\alpha_{\mathfrak{p},1},\alpha_{\mathfrak{p},2}}$ is an ordinary, unitary representation. By the classification of unitary representations (see e.g. [Ge75], Thm. 4.27), a spherical representation $\pi_{\alpha_{\mathfrak{p},1},\alpha_{\mathfrak{p},2}} = \pi(\chi_1, \chi_2)$ is unitary if and only if either χ_1, χ_2 are both unitary characters (i.e. $|\alpha_{\mathfrak{p},1}| = |\alpha_{\mathfrak{p},2}| = \sqrt{q_\mathfrak{p}}$), or $\chi_{1,2} = \chi_0|\cdot|^{\pm s}$ with χ_0 unitary and $-\frac{1}{2} < s < \frac{1}{2}$. A special representation $\pi_{\alpha_{\mathfrak{p},1},\alpha_{\mathfrak{p},2}} = \pi(\chi_1, \chi_2)$ is unitary if and only if the central character $\chi_1\chi_2$ is unitary. In all three cases, we have thus $\max\{|\alpha_{\mathfrak{p},1}|, |\alpha_{\mathfrak{p},2}|\} \geq \sqrt{q_\mathfrak{p}}$. Without loss of generality, we will assume the $\alpha_{\mathfrak{p},i}$ to be ordered such that $|\alpha_{\mathfrak{p},1}| \leq |\alpha_{\mathfrak{p},2}|$ for all $\mathfrak{p}|p$.
As in chapter 2, we define $a_\mathfrak{p} := \alpha_{\mathfrak{p},1} + \alpha_{\mathfrak{p},2}$, $\nu_\mathfrak{p} := \alpha_{\mathfrak{p},1}\alpha_{\mathfrak{p},2}/q_\mathfrak{p}$.
Let $\underline{\alpha_i} := (\alpha_{\mathfrak{p},i}, \mathfrak{p}|p)$, for $i = 1, 2$. We denote by $\mathfrak{A}_0(G, \underline{2}, \chi_Z, \underline{\alpha_1}, \underline{\alpha_2})$ the subset of all $\pi \in \mathfrak{A}_0(G, \underline{2}, \chi_Z)$ such that $\pi_\mathfrak{p} = \pi_{\alpha_{\mathfrak{p},1},\alpha_{\mathfrak{p},2}}$ for all $\mathfrak{p}|p$.
For later use we note that $\pi^\infty = \otimes_{v\nmid\infty}\pi_v$ is known to be defined over a finite extension of $\mathbb{Q}$, the smallest such field being the *field of definition* of π (cf. [Sp14]).

4.1 Upper half-space

For $k \in \{\mathbb{R}, \mathbb{C}\}$, let $\mathcal{H}_m := \mathcal{H}_k := k \times \mathbb{R}_+^*$ be the upper half-space of dimension $m := [k : \mathbb{R}] + 1$. Each $\mathcal{H}_m$ is a differentiable manifold of dimension m. If we write $x = (u, t) \in \mathcal{H}_m$ with $t \in \mathbb{R}_+^*$, u in $\mathbb{R}$ or $\mathbb{C}$, respectively, it has a Riemannian metric $ds^2 = \frac{dt^2 + du\, d\overline{u}}{t}$, which induces a hyperbolic geometry on $\mathcal{H}_m$, i.e. the geodesic lines on $\mathcal{H}_m$ are given by "vertical" lines $\{u\} \times \mathbb{R}_+^*$ and half-circles with center in the line or plane $t = 0$. $\mathcal{H}_\mathbb{R}$ is naturally isomorphic to the complex upper half-plane $\{z \in \mathbb{C}|\operatorname{Im}(z) > 0\}$.
We have the decompositions $\mathrm{GL}_2(\mathbb{C})^+ = B'_\mathbb{C} \cdot Z(\mathbb{C}) \cdot K_\mathbb{C}$ and $\mathrm{GL}_2(\mathbb{R})^+ = B'_\mathbb{R} \cdot Z(\mathbb{R}) \cdot K_\mathbb{R}$, where $B'_k \subseteq GL_2(k)$ is the subgroup of matrices $\left(\begin{smallmatrix} \mathbb{R}_+^* & k \\ 0 & 1 \end{smallmatrix}\right)$ for $k = \mathbb{R}, \mathbb{C}$, Z is the center, and $K_\mathbb{R} = \mathrm{SO}(2)$, $K_\mathbb{C} = \mathrm{SU}(2)$ (cf. [By98], Cor. 43). Identifying B'_k with $\mathcal{H}_k$ via $\left(\begin{smallmatrix} t & z \\ 0 & 1 \end{smallmatrix}\right) \mapsto (z, t)$ gives natural projections

$$\pi_\mathbb{R} : \mathrm{GL}_2(\mathbb{R})^+ \twoheadrightarrow \mathrm{GL}_2(\mathbb{R})^+/Z(\mathbb{R})\,\mathrm{SO}(2) \cong \mathcal{H}_2,$$

$$\pi_\mathbb{C} : \mathrm{GL}_2(\mathbb{C}) \twoheadrightarrow \mathrm{GL}_2(\mathbb{C})/Z(\mathbb{C})K_\mathbb{C} \cong \mathcal{H}_\mathbb{C}$$

and corresponding left $\mathrm{GL}_2(k)$-actions on cosets.
A differential form ω on $\mathcal{H}_m$ is called *left-invariant* if it is invariant under the pullback L_g^* of left multiplication $L_g : x \mapsto gx$ on $\mathcal{H}_m$, for all $g \in G$.

Following [By98], eqs. (4.20), (4.24), we choose the following basis of left-invariant differential 1-forms on $\mathcal{H}_3$:

$$\beta_0 := -\frac{dz}{t}, \quad \beta_1 := \frac{dt}{t}, \quad \beta_2 := \frac{d\overline{z}}{t},$$

and on $\mathcal{H}_2$ (writing $z = x + iy \in \mathcal{H}_2 \subseteq \mathbb{C}$):

$$\beta_1 := \frac{dz}{y}, \quad \beta_2 := -\frac{d\overline{z}}{y}.$$

We note that a form $f_1\beta_1 + f_2\beta_2$ is harmonic on $\mathcal{H}_2$ if and only if f_1/y and f_2/y are holomorphic functions in z ([By98], lemma 60).
The Jacobian $J(g,(0,1))$ of left multiplication by g in $(0,1) \in \mathcal{H}_m$ with respect to the basis $(\beta_i)_i$ gives rise to a representation

$$\varrho = \varrho_k : Z(k) \cdot K_k \to \mathrm{SL}_m(\mathbb{C})$$

with $\varrho|_{Z(k)}$ trivial, which on K_k is explicitly given by

$$\varrho_{\mathbb{C}}(h) = \begin{pmatrix} u^2 & 2uv & v^2 \\ -u\overline{v} & u\overline{u} - v\overline{v} & v\overline{u} \\ \overline{v}^2 & -2\overline{uv} & \overline{u}^2 \end{pmatrix} \quad \text{for } h = \begin{pmatrix} u & v \\ -\overline{v} & \overline{u} \end{pmatrix} \in \mathrm{SU}(2),$$

resp.

$$\varrho_{\mathbb{R}} \begin{pmatrix} \cos(\vartheta) & \sin(\vartheta) \\ -\sin(\vartheta) & \cos(\vartheta) \end{pmatrix} = \begin{pmatrix} e^{2i\vartheta} & 0 \\ 0 & e^{-2i\vartheta} \end{pmatrix}$$

([By98], (4.27), (4.21)). In the real case, we will only consider harmonic forms on $\mathcal{H}_2$ that are multiples of β_1, thus we sometimes identify $\varrho_{\mathbb{R}}$ with its restriction $\varrho_{\mathbb{R}}^{(1)}$ to the first basis vector β_1,

$$\varrho_{\mathbb{R}}^{(1)} : \mathrm{SO}(2) \to S^1 \subseteq \mathbb{C}^*, \quad \kappa_\vartheta = \begin{pmatrix} \cos(\vartheta) & \sin(\vartheta) \\ -\sin(\vartheta) & \cos(\vartheta) \end{pmatrix} \mapsto e^{2i\vartheta}.$$

For each i, let ω_i be the left-invariant differential 1-form on $\mathrm{GL}_2(k)$ which coincides with the pullback $(\pi_{\mathbb{C}})^*\beta_i$ at the identity. Write $\underline{\omega}$ (resp. $\underline{\beta}$) for the column vector of the ω_i (resp. β_i). Then we have the following lemma from [By98]:

LEMMA 4.2. *For each i, the differential ω_i on G induces β_i on $\mathcal{H}_m$, by restriction to the subgroup $B'_k \cong \mathcal{H}_m$. For a function $\phi : G \to \mathbb{C}^m$, the form $\phi \cdot \underline{\omega}$ (with $\mathbb{C}^m$ considered as a row vector, so $\cdot$ is the scalar product of vectors) induces $f \cdot \underline{\beta}$, where $f : \mathcal{H}_m \to \mathbb{C}^m$ is given by*

$$f(z,t) := \phi\left(\begin{pmatrix} t & z \\ 0 & 1 \end{pmatrix}\right).$$

(See [By98], Lemma 57.)
To consider the infinite primes of F all at once, we define

$$\mathcal{H}_\infty := \prod_{i=0}^{d} \mathcal{H}_{m_i} = \prod_{i=0}^{r-1} \mathcal{H}_2 \times \prod_{i=r}^{d} \mathcal{H}_3$$

(where $m_i = 2$ if σ_i is a real embedding, and $m_i = 3$ if σ_i is complex), and let $\mathcal{H}_\infty^0 := \prod_{i=1}^d \mathcal{H}_{m_i}$ be the product with the zeroth factor removed. (The choice of the 0-th factor is for convenience; we could also choose any other infinite place, whether real or complex.)

For each embedding σ_i, the elements of $\mathbb{P}^1(F)$ are cusps of $\mathcal{H}_{m_i}$: for a given complex embedding $F \hookrightarrow \mathbb{C}$, we can identify F with $F \times \{0\} \hookrightarrow \mathbb{C} \times \mathbb{R}_{\geq 0}$ and define the "extended upper half-space" as $\overline{\mathcal{H}_3} := \mathcal{H}_3 \cup F \cup \{\infty\} \subseteq \mathbb{C} \times \mathbb{R}_{\geq 0} \cup \{\infty\}$; similarly for a given real embedding $F \hookrightarrow \mathbb{R}$, we get the extended upper half-plane $\overline{\mathcal{H}_2} := \mathcal{H}_2 \cup F \cup \{\infty\}$. A basis of neighbourhoods of the cusp ∞ is given by the sets $\{(u,t) \in \mathcal{H}_m | t > N\}$, $N \gg 0$, and of $x \in F$ by the open half-balls in $\mathcal{H}_m$ with center $(x,0)$.
Let $G(F)^+ \subseteq G(F)$ denote the subgroup of matrices with totally positive determinant. It acts on $\mathcal{H}_\infty^0$ by composing the embedding

$$G(F)^+ \hookrightarrow \prod_{v|\infty, v \neq v_0} G(F_v)^+, \qquad g \mapsto (\sigma_1(g), \ldots, \sigma_d(g)),$$

with the actions of $G(\mathbb{C})^+ = G(\mathbb{C})$ on $\mathcal{H}_3$ and $G(\mathbb{R})^+$ on $\mathcal{H}_2$ as defined above, and on $\Omega^d_{\text{harm}}(\mathcal{H}_\infty^0)$ by the inverse of the corresponding pullback, $\gamma \cdot \underline{\omega} := (\gamma^{-1})^* \underline{\omega}$. Both are left actions.
For each complex v, we write the codomain of ϱ_{F_v} as

$$\varrho_{F_v} : Z(F_v) \cdot K_{F_v} \to \mathrm{SL}_3(\mathbb{C}) =: \mathrm{SL}(V_v),$$

for a three-dimensional $\mathbb{C}$-vector space V_v. We denote the harmonic forms on $\mathrm{GL}_2(F_v)$, $\mathcal{H}_{F_v}$ defined above by $\underline{\omega_v}$, $\underline{\beta_v}$ etc.
Let $V = \bigotimes_{v \in S_\mathbb{C}} V_v \cong (\mathbb{C}^3)^{\otimes s}$, $Z_\infty = \prod_{v|\infty} Z(F_v)$, $K_\infty = \prod_{v|\infty} K_{F_v}$. Denoting by $S_\mathbb{C}$ (resp. $S_\mathbb{R}$) the set of complex (resp. real) archimedian primes of F, we can merge the representations ϱ_{F_v} for each $v|\infty$ into a representation

$$\varrho = \varrho_\infty := \bigotimes_{v \in S_\mathbb{C}} \varrho_\mathbb{C} \otimes \bigotimes_{v \in S_\mathbb{R}} \varrho_\mathbb{R}^{(1)} : Z_\infty \cdot K_\infty \to \mathrm{SL}(V),$$

and define V-valued vectors of differential forms

$$\underline{\omega} := \bigotimes_{v \in S_\mathbb{C}} \underline{\omega_v} \otimes \bigotimes_{v \in S_\mathbb{R}} \omega_v^1, \quad \underline{\beta} := \bigotimes_{v \in S_\mathbb{C}} \underline{\beta_v} \otimes \bigotimes_{v \in S_\mathbb{R}} (\beta_v)_1$$

on $\mathrm{GL}_2(F_\infty)$ and $\mathcal{H}_\infty$, respectively.

4.2 Automorphic forms

Let $\chi_Z : \mathbb{A}_F^*/F^* \to \mathbb{C}^*$ be a Hecke character that is trivial at the archimedian places. We also denote by χ_Z the corresponding character on $Z(\mathbb{A}_F)$ under the isomorphism $\mathbb{A}_F^* \to Z(\mathbb{A}_F)$, $a \mapsto \left(\begin{smallmatrix} a & 0 \\ 0 & a \end{smallmatrix}\right)$.

Definition 4.3. An *automorphic cusp form of parallel weight $\underline{2}$ with central character χ_Z* is a map $\phi : G(\mathbb{A}_F) \to V$ such that

(i) $\phi(z\gamma g) = \chi_Z(z)\phi(g)$ for all $g \in G(\mathbb{A})$, $z \in Z(\mathbb{A})$, $\gamma \in G(F)$.

(ii) $\phi(gk_\infty) = \phi(g)\varrho(k_\infty)$ for all $k_\infty \in K_\infty$, $g \in G(\mathbb{A})$ (considering V as a row vector).

(iii) ϕ has "moderate growth" on $B'_{\mathbb{A}} := \{\begin{pmatrix} y & x \\ 0 & 1 \end{pmatrix} \in G(\mathbb{A})\}$, i.e. $\exists C, \lambda\ \forall A \in B'_{\mathbb{A}} : \|\phi(A)\| \leq C \cdot \sup(|y|^\lambda, |y|^{-\lambda})$ (for any fixed norm $\|\cdot\|$ on V);
and $\phi|_{G(\mathbb{A}_\infty)} \cdot \underline{\omega}$ is the pullback of a harmon ic form $\omega_\phi = f_\phi \cdot \underline{\beta}$ on $\mathcal{H}_\infty$.

(iv) There exists a compact open subgroup $K' \subseteq G(\mathbb{A}^\infty)$ such that $\phi(gk) = \phi(g)$ for all $g \in G(\mathbb{A})$ and $k \in K'$.

(v) For all $g \in G(\mathbb{A}_F)$,

$$\int_{\mathbb{A}_F/F} \phi\left(\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} g\right)\, dx = 0. \qquad \textit{("Cuspidality")}$$

We denote by $\mathcal{A}_0(G, \mathrm{harm}, \underline{2}, \chi_Z)$ the space of all such maps ϕ.

For each $g^\infty \in \mathbb{A}_F^\infty$, let $\omega_\phi(g^\infty)$ be the restriction of $\phi(g^\infty, \cdot) \cdot \underline{\omega}$ from $G(\mathbb{A}_F^\infty)$ to $\mathcal{H}_\infty$; it is a $(d+1)$-form on $\mathcal{H}_\infty$.
We want to integrate $\omega_\phi(g^\infty)$ between two cusps of the space $\mathcal{H}_{m_0}$. (We will identify each $x \in \mathbb{P}^1(F)$ with its corresponding cusp in $\overline{\mathcal{H}_{m_0}}$ in the following.) The geodesic between the cusps $x \in F$ and ∞ in $\overline{\mathcal{H}_{m_0}}$ is the line $\{x\} \times \mathbb{R}_+^* \subseteq \mathcal{H}_{m_0}$ and the integral of ω_ϕ along it is finite since ϕ is uniformly rapidly decreasing:

Theorem 4.4. *(Gelfand, Piatetski-Shapiro) An automorphic cusp form ϕ is rapidly decreasing modulo the center on a fundamental domain $\mathcal{F}$ of $\mathrm{GL}_2(F)\backslash \mathrm{GL}_2(\mathbb{A}_F)$;*
i.e. there exists an integer r such that for all $N \in \mathbb{N}$ there exists a $C > 0$ such that

$$\phi(zg) \leq C|z|^r \|g\|^{-N}$$

for all $z \in Z(\mathbb{A}_F)$, $g \in \mathcal{F} \cap \mathrm{SL}_2(\mathbb{A}_F)$. Here $\|g\| := \max\{|g_{i,j}|, |(g^{-1})_{i,j}|\}_{i,j \in \{1,2\}}$.

(See [CKM04], Thm. 2.2; or [Kur78], (6) for quadratic imaginary F.)
In fact, the integral of $\omega_\phi(g^\infty)$ along $\{x\} \times \mathbb{R}_+^* \subseteq \mathcal{H}_{m_0}$ equals the integral of $\phi(g^\infty, \cdot) \cdot \underline{\omega}$ along a path $g_t \in \mathrm{GL}_2(F_{\infty_0})$, $t \in \mathbb{R}_+^*$, where we can choose

$$g_t = \frac{1}{\sqrt{t}} \begin{pmatrix} t & x \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{t}} & \frac{x}{\sqrt{t}} \\ 0 & \sqrt{t} \end{pmatrix},$$

and thus have $\|g_t\| = \sqrt{t}$ for all $t \gg 0$, $\|g_t\| = C\frac{1}{\sqrt{t}}$ for $t \ll 1$, so the integral $\int_x^\infty \omega_\phi(g^\infty) \in \Omega^d_{\mathrm{harm}}(\mathcal{H}^0_\infty)$ is well-defined by the theorem.
For any two cusps $a, b \in \mathbb{P}^1(F)$, we now define

$$\int_a^b \omega_\phi(g^\infty) := \int_a^\infty \omega_\phi(g^\infty) - \int_b^\infty \omega_\phi(g^\infty) \in \Omega^d_{\mathrm{harm}}(\mathcal{H}^0_\infty).$$

Since ϕ is *uniformly* rapidly decreasing ($\|g_t\|$ does not depend on x, for $t \gg 0$), this integral along the path $(a, 0) \to (a, \infty) = (b, \infty) \to (b, 0)$ in $\overline{\mathcal{H}}_{m_0}$ is the same as the limit (for $t \to \infty$) of the integral along $(a, 0) \to (a, t) \to (b, t) \to (b, 0)$; and since ω_ϕ is harmonic (and thus integration is path-independent within $\mathcal{H}_{m_0}$) the latter is in fact independent of t, so equality holds for each $t > 0$, or along any path from $(a, 0)$ to $(b, 0)$ in $\mathcal{H}_{m_0}$. Thus $\int_a^b \omega_\phi(g^\infty)$ equals the integral of $\omega_\phi(g^\infty)$ along the geodesic from a to b, and we have

$$\int_a^b \omega_\phi(g^\infty) + \int_b^c \omega_\phi(g^\infty) = \int_a^c \omega_\phi(g^\infty)$$

for any three cusps $a, b, c \in \mathbb{P}^1(F)$. Let $\mathrm{Div}(\mathbb{P}^1(F))$ denote the free abelian group of divisors of $\mathbb{P}^1(F)$, and let $\mathcal{M} := \mathrm{Div}_0(\mathbb{P}^1(F))$ be the subgroup of divisors of degree 0.
We can extend the definition of the integral linearly to get a homomorphism

$$\mathcal{M} \to \Omega^d_{\mathrm{harm}}(\mathcal{H}^0_\infty), \quad m \mapsto \int_m \omega_\phi(g^\infty),$$

and easily check that

$$\gamma^* \left(\int_{\gamma m} \omega_\phi(\gamma g) \right) = \int_m \omega_\phi(g). \tag{29}$$

for all $\gamma \in G(F)^+$, $g \in G(\mathbb{A}^\infty)$, $m \in \mathcal{M}$.
Now let $\mathfrak{m}$ be an ideal of F prime to p, let χ_Z be a Hecke character of conductor dividing $\mathfrak{m}$, and $\underline{\alpha_1}, \underline{\alpha_2}$ as above.

Definition 4.5. We define $S_2(G, \mathfrak{m}, \underline{\alpha_1}, \underline{\alpha_2})$ to be the $\mathbb{C}$-vector space of all maps

$$\Phi : G(\mathbb{A}^p) \to \mathcal{B}^{\underline{\alpha_1}, \underline{\alpha_2}}(F_p, V) = \mathrm{Hom}(\mathcal{B}_{\underline{\alpha_1}, \underline{\alpha_2}}(F_p, \mathbb{C}), V)$$

such that:

(a) ϕ is "almost" $K_0(\mathfrak{m})$-invariant (in the notation of [Ge75]), i.e. $\phi(gk) = \phi(g)$ for all $g \in G(\mathbb{A}^p)$ and $k \in \prod_{v \nmid \mathfrak{m}p} G(\mathcal{O}_v)$, and $\phi(gk) = \chi_Z(a)\phi(g)$ for all $v|\mathfrak{m}$, $k = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in K_0(\mathfrak{m})_v$ and $g \in G(\mathbb{A}^p)$.

(b) For each $\psi \in \mathcal{B}_{\underline{\alpha_1},\underline{\alpha_2}}(F_p, \mathbb{C})$, the map

$$\langle \Phi, \psi \rangle : G(\mathbb{A}) = G(F_p) \times G(\mathbb{A}^p) \to V,\ (g_p, g^p) \mapsto \Phi(g^p)(g_p\psi)$$

lies in $\mathcal{A}_0(G, \mathrm{harm}, \underline{2}, \chi_Z)$.

Note that (a) implies that ϕ is K'-invariant for some open subgroup $K' \subseteq K_0(\mathfrak{m})^p$ of finite index ([By98]/[We71]).

4.3 Cohomology of $\mathrm{GL}_2(F)$

Let M be a left $G(F)$-module and N an $R[H]$-module, for a ring R and a subgroup $H \subseteq G(F)$. Let $S \subseteq S_p$ be a set of primes of F dividing p; as above, let $\chi = \chi_Z$ be a Hecke character of conductor $\mathfrak{m}$ prime to p.

Definition 4.6. For a compact open subgroup $K \subseteq K_0(\mathfrak{m})^S \subseteq G(\mathbb{A}^{S,\infty})$, we denote by $\mathcal{A}_f(K, S, M; N)$ the R-module of all maps $\Phi : G(\mathbb{A}^{S,\infty}) \times M \to N$ such that

1. $\Phi(gk, m) = \Phi(g, m)$ for all $g \in G(\mathbb{A}^{S,\infty})$, $m \in M$, $k \in \prod_{v \nmid \mathfrak{m}p} G(\mathcal{O}_v)$;
2. $\Phi(gk) = \chi_Z(a)\Phi(g)$ for all $v|\mathfrak{m}$, $k = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in K_0(\mathfrak{m})_v$ and $g \in G(\mathbb{A}^{S,\infty})$, $m \in M$.

We denote by $\mathcal{A}_f(S, M; N)$ the union of the $\mathcal{A}_f(K, S, M; N)$ over all compact open subgroups K.

$\mathcal{A}_f(S, M; N)$ is a left $G(\mathbb{A}^{S,\infty})$-module via $(\gamma \cdot \Phi)(g, m) := \Phi(\gamma^{-1}g, m)$ and has a left H-operation given by $(\gamma \cdot \Phi)(g, m) := \gamma\Phi(\gamma^{-1}g, \gamma^{-1}m)$, commuting with the $G(\mathbb{A}^{S,\infty})$-operation.
In contrast to our previous notation, we consider two subsets $S_1 \subseteq S_2 \subseteq S_p$ in this section. We put $(\underline{\alpha_1}, \underline{\alpha_2})_{S_1} := \{(\alpha_{\mathfrak{p},1}, \alpha_{\mathfrak{p},2}) | \mathfrak{p} \in S_1\}$, we set

$$\mathcal{A}_f((\underline{\alpha_1}, \underline{\alpha_2})_{S_1}, S_2, M; N) = \mathcal{A}_f(S_2, M; \mathcal{B}^{(\underline{\alpha_1},\underline{\alpha_2})_{S_1}}(F_{S_1}, N));$$

we write $\mathcal{A}_f(\mathfrak{m}, (\underline{\alpha_1}, \underline{\alpha_2})_{S_1}, S_2, M; N) := \mathcal{A}_f(K_0(\mathfrak{m}), (\underline{\alpha_1}, \underline{\alpha_2})_{S_1}, S_2, M; N)$. If $S_1 = S_2$, we will usually drop S_2 from all these notations.
We have a natural identification of $\mathcal{A}_f(\mathfrak{m}, (\underline{\alpha_1}, \underline{\alpha_2})_S, M; N)$ with the space of maps $G(\mathbb{A}^{S,\infty}) \times M \times \mathcal{B}_{(\underline{\alpha_1},\underline{\alpha_2})_S}(F_S, R) \to N$ that are "almost" K-invariant.
Let $S_0 \subseteq S_1 \subseteq S_2 \subseteq S_p$ be subsets. The pairing (11) induces a pairing

$$\mathcal{A}_f((\underline{\alpha_1}, \underline{\alpha_2})_{S_1}, S_2, M; N) \times \mathcal{B}_{(\underline{\alpha_1},\underline{\alpha_2})_{S_0}}(F_{S_0}, R) \to \mathcal{A}_f((\underline{\alpha_1}, \underline{\alpha_2})_{S_0}, S_2, M; N)$$

which, when restricting to K-invariant elements, induces an isomorphism

$$\mathcal{A}_f(K,(\underline{\alpha_1},\underline{\alpha_2})_{S_1},S_2,M;N)\cong\mathcal{B}^{(\underline{\alpha_1},\underline{\alpha_2})_{S_1-S_0}}(F_{S_1-S_0},\mathcal{A}_f(\underline{\alpha_1},\underline{\alpha_2})_{S_0},S_2,M;N).$$

Putting $S_0:=S_1-\{\mathfrak{p}\}$ for a prime $\mathfrak{p}\in S_1$, we specifically get an isomorphism

$$\mathcal{A}_f(K,(\underline{\alpha_1},\underline{\alpha_2})_{S_1},S_2,M;N)\cong\mathcal{B}^{\alpha_{\mathfrak{p},1},\alpha_{\mathfrak{p},2}}(F_{\mathfrak{p}},\mathcal{A}_f(\underline{\alpha_1},\underline{\alpha_2})_{S_0},S_2,M;N).$$

Lemmas 2.9 and 2.10 now immediately imply the following:

Lemma 4.7. *Let $S\subseteq S_p$, $\mathfrak{p}\in S$, $S_0:=S-\{\mathfrak{p}\}$. Let $K\subseteq G(\mathbb{A}^{S,\infty})$ be a compact open subgroup.*
(a) If $\pi_{\alpha_{\mathfrak{p},1},\alpha_{\mathfrak{p},2}}$ is spherical, we have exact sequences

$$0\to\mathcal{A}_f(K,(\underline{\alpha_1},\underline{\alpha_2})_S,M;N)\to Z\xrightarrow{N-\nu_{\mathfrak{p}}}Z\to 0$$

and

$$0\to Z\to\mathcal{A}_f(K_0,(\underline{\alpha_1},\underline{\alpha_2})_{S_0},M;N)\xrightarrow{T-a_{\mathfrak{p}}}\mathcal{A}_f(K_0,(\underline{\alpha_1},\underline{\alpha_2})_{S_0},M;N)\to 0$$

for a $G(\mathbb{A}^{S_0,\infty})$-module Z and a compact open subgroup $K_0=K\times K_{\mathfrak{p}}$ of $G(\mathbb{A}^{S_0,\infty})$.
(b) If $\pi_{\alpha_{\mathfrak{p},1},\alpha_{\mathfrak{p},2}}$ is special (with central character $\chi_{\mathfrak{p}}$), we have exact sequences

$$0\to\mathcal{A}_f(K,(\underline{\alpha_1},\underline{\alpha_2})_S,M;N)\to Z'\to Z\to 0$$

and

$$\begin{aligned}0\to&Z\to\mathcal{A}_f(K_0,(\underline{\alpha_1},\underline{\alpha_2})_{S_0},M;N)^2\to\mathcal{A}_f(K_0,(\underline{\alpha_1},\underline{\alpha_2})_{S_0},M;N)^2\to 0,\\0\to&Z'\to\mathcal{A}_f(K_0',(\underline{\alpha_1},\underline{\alpha_2})_{S_0},M;N)^2\to\mathcal{A}_f(K_0',(\underline{\alpha_1},\underline{\alpha_2})_{S_0},M;N)^2\to 0,\end{aligned}$$

with $Z^{(')}:=\mathcal{A}_f(K_0^{(')},(\underline{\alpha_1},\underline{\alpha_2})_{S_0},S,M;N(\chi_{\mathfrak{p}}))$, where $K_0^{(')}=K\times K_{\mathfrak{p}}^{(')}$ are compact open subgroups of $G(\mathbb{A}^{S_0,\infty})$.

Proposition 4.8. *Let $S\subseteq S_p$ and let K be a compact open subgroup of $G(\mathbb{A}^{S,\infty})$.*
(a) For each flat R-module N (with trivial $G(F)$-action), the canonical map

$$H^q(G(F)^+,\mathcal{A}_f(K,(\underline{\alpha_1},\underline{\alpha_2})_S,\mathcal{M};R))\otimes_R N\to H^q(G(F)^+,\mathcal{A}_f(K,(\underline{\alpha_1},\underline{\alpha_2})_S,\mathcal{M};N))$$

is an isomorphism for each $q\geq 0$.
(b) If R is finitely generated as a $\mathbb{Z}$-module, $H^q(G(F)^+,\mathcal{A}_f(K,(\underline{\alpha_1},\underline{\alpha_2})_S,\mathcal{M};R)$ is finitely generated over R.

Proof. We can copy the proof of [Sp14], Prop. 5.6, using lemma 4.7 instead of [Sp14], lemma 5.4 to reduce to the case $S=\varnothing$. ∎

We define

$$H^q_*(G(F)^+, \mathcal{A}_f((\underline{\alpha_1}, \underline{\alpha_2})_S, M; R)) := \varinjlim H^q(G(F)^+, \mathcal{A}_f(K, (\underline{\alpha_1}, \underline{\alpha_2})_S, M; R))$$

where the limit runs over all compact open subgroups $K \subseteq G(\mathbb{A}^{S,\infty})$; and similarly define $H^q_*(B(F)^+, \mathcal{A}_f((\underline{\alpha_1}, \underline{\alpha_2})_S, \mathcal{M}; R)$. The proposition immediately implies

Corollary 4.9. *Let $R \to R'$ be a flat ring homomorphism. Then the canonical map*

$$H^q_*(G(F)^+, \mathcal{A}_f((\underline{\alpha_1}, \underline{\alpha_2})_S, \mathcal{M}; R)) \otimes_R R' \to H^q_*(G(F)^+, \mathcal{A}_f((\underline{\alpha_1}, \underline{\alpha_2})_S, \mathcal{M}; R')$$

is an isomorphism, for all $q \geq 0$.

If $R = k$ is a field of characteristic zero, $H^q_*(G(F)^+, \mathcal{A}_f((\underline{\alpha_1}, \underline{\alpha_2})_S, M; R)$ is a smooth $G(\mathbb{A}^{S,\infty})$-module, and we have

$$H^q_*(G(F)^+, \mathcal{A}_f((\underline{\alpha_1}, \underline{\alpha_2})_S, M; k)^K = H^q(G(F)^+, \mathcal{A}_f(K, (\underline{\alpha_1}, \underline{\alpha_2})_S, M; k).$$

We identify $G(F)/G(F)^+$ with the group $\Sigma = \{\pm 1\}^r$ via the isomorphism

$$G(F)/G(F^+) \xrightarrow{\det} F^*/F^*_+ \cong \Sigma$$

(with all groups being trivial for $r = 0$). Then Σ acts on $H^q_*(G(F)^+, \mathcal{A}_f((\underline{\alpha_1}, \underline{\alpha_2})_S, M; k)$ and $H^q(G(F)^+, \mathcal{A}_f(K, (\underline{\alpha_1}, \underline{\alpha_2})_S, M; k)$ by conjugation.For $\pi \in \mathfrak{A}_0(G, \underline{2})$ and $\underline{\mu} \in \Sigma$, we write $H^q_*(G(F)^+, \cdot)_{\pi,\underline{\mu}} := \mathrm{Hom}_{G(\mathbb{A}^{S,\infty})}(\pi^S, H^q_*(G(F)^+, \cdot))_{\underline{\mu}}$.

Proposition 4.10. *Let $\pi \in \mathfrak{A}_0(G, \underline{2}, \chi_Z, \underline{\alpha_1}, \underline{\alpha_2})$, $S \subseteq S_p$. Let k be a field which contains the field of definition of π. Then for every $\underline{\mu} \in \Sigma$, we have*

$$H^q_*(G(F)^+, \mathcal{A}_f((\underline{\alpha_1}, \underline{\alpha_2})_S, \mathcal{M}; k)_{\pi,\underline{\mu}} = \begin{cases} k, & \textit{if } q = d; \\ 0, & \textit{if } q \in \{0, \ldots, d-1\} \end{cases} \tag{30}$$

Proof. The case $S = \varnothing$ is proved analogously to [Sp14], prop. 5.8, using the results of Harder [Ha87]. For $S = S_0 \cup \{\mathfrak{p}\}$ and $\pi_\mathfrak{p}$ spherical, lemma 4.7(a) and the statement for S_0 give an isomorphism

$$H^q_*(G(F)^+, \mathcal{A}_f((\underline{\alpha_1}, \underline{\alpha_2})_{S_0}, \mathcal{M}; k))_{\pi,\underline{\mu}} \cong H^q_*(G(F)^+, \mathcal{A}_f((\underline{\alpha_1}, \underline{\alpha_2})_S, \mathcal{M}; k))_{\pi,\underline{\mu}}$$

since the Hecke operators $T_\mathfrak{p}$ and $N_\mathfrak{p}$ act on the left-hand side by multiplication with $a_\mathfrak{p}$ and $\nu_\mathfrak{p}$, respectively. If $\pi_\mathfrak{p}$ is special, we can similarly deduce the statement for S from that for S_0, using the first exact sequence of lemma 4.7(b), since the results of [Ha87] also hold when twisting k by a (central) character. $\square$

4.4 Eichler-Shimura map

From now on, let $S_1 \subseteq S_p$ be the set of places such that $\pi_\mathfrak{p}$ is the Steinberg representation (i.e. $\alpha_{\mathfrak{p},1} = \nu_\mathfrak{p} = 1$, $\alpha_{\mathfrak{p},2} = q$).
Given a subgroup $K_0(\mathfrak{m})^p \subseteq G(\mathbb{A}^{p,\infty})$ as above, there is a map

$$I_0 : S_2(G, \mathfrak{m}, \underline{\alpha_1}, \underline{\alpha_2}) \to H^0(G(F)^+, \mathcal{A}_f(\mathfrak{m}, \underline{\alpha_1}, \underline{\alpha_2}, \mathcal{M}; \Omega^d_{\text{harm}}(\mathcal{H}^0_\infty)))$$

given by

$$I_0(\Phi) : (\psi, (g, m)) \mapsto \int_m \omega_{\langle\Phi,\psi\rangle}(1_p, g),$$

for $\psi \in \mathcal{B}_{\underline{\alpha_1},\underline{\alpha_2}}(F_p, \mathbb{C}), g \in G(\mathbb{A}^{p,\infty}), m \in \mathcal{M}$, where 1_p denotes the unity element in $G(F_p)$.
This is well-defined since both sides are "almost" $K_0(\mathfrak{m})$-invariant, and the $G(F)^+$-invariance of $I_0(\Phi)$ follows from a straightforward calculation, using (29).
From the complex

$$\mathcal{A}_f(m, \underline{\alpha_1}, \underline{\alpha_2}, \mathcal{M}; \mathbb{C}) \to C^\bullet := \mathcal{A}_f(\mathfrak{m}, \underline{\alpha_1}, \underline{\alpha_2}, \mathcal{M}; \Omega^\bullet_{\text{harm}}(\mathcal{H}^0_\infty))$$

we get a map

$$S_2(G, \mathfrak{m}, \underline{\alpha_1}, \underline{\alpha_2}) \to H^d(G(F)^+, \mathcal{A}_f(\mathfrak{m}, \underline{\alpha_1}, \underline{\alpha_2}, \mathcal{M}; \mathbb{C})) \tag{31}$$

by composing I_0 with an edge morphism of the spectral sequence

$$H^q(G(F)^+, C^p) \implies H^{p+q}(G(F)^+, C^\bullet).$$

Using the map $\delta^{\underline{\alpha_1},\underline{\alpha_2}} : \mathcal{B}^{\underline{\alpha_1},\underline{\alpha_2}}(F, V) \to \text{Dist}(F_p^*, V)$ from section 2.6, we next define a map

$$\Delta_V^{\underline{\alpha_1},\underline{\alpha_2}} : S_2(G, \mathfrak{m}, \underline{\alpha_1}, \underline{\alpha_2}) \to \mathcal{D}(S_1, V) \tag{32}$$

by

$$\Delta_V^{\underline{\alpha_1},\underline{\alpha_2}}(\Phi)(U, x^p) = \delta^{\underline{\alpha_1},\underline{\alpha_2}}\left(\Phi\begin{pmatrix} x^p & 0 \\ 0 & 1\end{pmatrix}\right)(U)$$

for $U \in \mathfrak{Co}(F_{S_1} \times F_{S_2}), x^p \in \mathbb{I}^p$, and we denote by $\Delta^{\underline{\alpha_1},\underline{\alpha_2}} : S_2(G, \mathfrak{m}, \underline{\alpha_1}, \underline{\alpha_2}) \to \mathcal{D}(S_1, \mathbb{C})$ its (1,...,1)th coordinate function (i.e. corresponding to the harmonic forms $\bigotimes_{v|\infty}(\omega_v)_1$, $\bigotimes_{v|\infty}(\beta_v)_1$ in section 4.1):

$$\Delta^{\underline{\alpha_1},\underline{\alpha_2}}(\Phi)(U, x^p) = \delta^{\underline{\alpha_1},\underline{\alpha_2}}\left(\Phi\begin{pmatrix} x^p & 0 \\ 0 & 1\end{pmatrix}\right)_{(1,\ldots,1)}(U).$$

Since for each complex prime v, $S^1 \cong \text{SU}(2) \cap T(\mathbb{C})$ operates on Φ via ϱ_v, $\Delta^{\underline{\alpha_1},\underline{\alpha_2}}$ is easily seen to be S^1-invariant, i.e. it lies in $\mathcal{D}'(S_1, \mathbb{C})$.
We also have a natural (i.e. commuting with the complex maps of each $C^\bullet$) family of maps

$$\mathcal{A}_f(\mathfrak{m}, \underline{\alpha_1}, \underline{\alpha_2}, \mathcal{M}, \Omega^i_{\text{harm}}(\mathcal{H}^0_\infty)) \to \mathcal{D}_f(S_1, \Omega^i(U^0_\infty, \mathbb{C})) \tag{33}$$

for all $i \geq 0$, and

$$\mathcal{A}_f(\mathfrak{m}, \underline{\alpha_1}, \underline{\alpha_2}, \mathcal{M}, \mathbb{C}) \to \mathcal{D}_f(S_1, \mathbb{C}) \tag{34}$$

(the $i = -1$-th term in the complexes), by mapping $\Phi \in \mathcal{A}_f(\mathfrak{m}, \underline{\alpha_1}, \underline{\alpha_2}, \mathcal{M}, \cdot)$ first to

$$(U, x^{p,\infty}) \mapsto \Phi\left(\begin{pmatrix} x^{p,\infty} & 0 \\ 0 & 1 \end{pmatrix}, \infty - 0\right)(\delta_{\underline{\alpha_1}, \underline{\alpha_2}}(1_U)) \in \Omega^i_{\mathrm{harm}}(\mathcal{H}^0_\infty) \text{ resp. } \in \mathbb{C},$$

and then for $i \geq 0$ restricting the differential forms to $\Omega^i(U^0_\infty)$ via

$$U^0_\infty = \prod_{v \in S^0_\infty} \mathbb{R}^*_+ \hookrightarrow \prod_{v \in S^0_\infty} \mathcal{H}_v = \mathcal{H}^0_\infty.$$

One easily checks that (33) and (34) are compatible with the homomorphism of "acting groups" $F^{*\prime} \hookrightarrow G(F)^+, x \mapsto \left(\begin{smallmatrix} x & 0 \\ 0 & 1 \end{smallmatrix}\right)$, so we get induced maps in cohomology

$$H^0(G(F)^+, \mathcal{A}_f(\mathfrak{m}, \underline{\alpha_1}, \underline{\alpha_2}, \mathcal{M}, \Omega^d_{\mathrm{harm}}(\mathcal{H}^0_\infty))) \to H^0(F^{*\prime}, \mathcal{D}_f(S_1, \Omega^d(U^0_\infty, \mathbb{C}))) \tag{35}$$

and

$$H^d(G(F)^+, \mathcal{A}_f(\mathfrak{m}, \underline{\alpha_1}, \underline{\alpha_2}, \mathcal{M}, \mathbb{C})) \to H^d(F^{*\prime}, \mathcal{D}_f(S_1, \mathbb{C})), \tag{36}$$

which are linked by edge morphisms of the respective spectral sequences to give a commutative diagram (given in the proof below).

PROPOSITION 4.11. *We have a commutative diagram:*

$$\begin{tikzcd}
S_2(G, \mathfrak{m}, \underline{\alpha_1}, \underline{\alpha_2}) \arrow[r, "(31)"] \arrow[d, "\Delta^{\underline{\alpha_1}, \underline{\alpha_2}}"] & H^d(G(F)^+, \mathcal{A}_f(\mathfrak{m}, \underline{\alpha_1}, \underline{\alpha_2}, \mathcal{M}, \mathbb{C})) \arrow[d, "(36)"] \\
\mathcal{D}'(S_1, \mathbb{C}) \arrow[r, "\phi \mapsto \kappa_\phi"] & H^d\big(F^{*\prime}, \mathcal{D}_f(S_1, \mathbb{C})\big)
\end{tikzcd}$$

Proof. The given diagram factorizes as

$$\begin{tikzcd}
S_2(G, \mathfrak{m}, \underline{\alpha_1}, \underline{\alpha_2}) \arrow[r, "I_0"] \arrow[d, "\Delta^{\underline{\alpha_1}, \underline{\alpha_2}}"] & H^0(G(F)^+, \mathcal{A}_f(\Omega^d_{\mathrm{harm}}(\mathcal{H}^0_\infty))) \arrow[r] \arrow[d, "(35)"] & H^d(G(F)^+, \mathcal{A}_f(\mathbb{C})) \arrow[d, "(36)"] \\
\mathcal{D}'(S_1, \mathbb{C}) \arrow[r] & H^0(F^{*\prime}, \mathcal{D}_f(S_1, \Omega^d(U^0_\infty, \mathbb{C}))) \arrow[r] & H^d\big(F^{*\prime}, \mathcal{D}_f(S_1, \mathbb{C})\big)
\end{tikzcd}$$

(where we write $\mathcal{A}_f(\cdot)$ instead of $\mathcal{A}_f(\mathfrak{m}, \underline{\alpha_1}, \underline{\alpha_2}, \mathcal{M}, \cdot)$ for brevity). The right-hand square is the naturally commutative square mentioned above; the commutativity of the left-hand square can easily be checked by hand. $\square$

4.5 Whittaker model

We now consider an automorphic representation $\pi = \otimes_\nu \pi_\nu \in \mathfrak{A}_0(G, \underline{2}, \chi_Z, \underline{\alpha_1}, \underline{\alpha_2})$. Denote by $\mathfrak{c}(\pi) := \prod_{v \text{ finite}} \mathfrak{c}(\pi_v)$ the conductor of π.

Let $\chi : \mathbb{I}^\infty \to \mathbb{C}^*$ be a unitary character of the finite ideles; for each finite place v, set $\chi_v = \chi|_{F_v^*}$. For each prime v of F, let $\mathcal{W}_v$ denote the Whittaker model of π_v. For each finite and each real prime, we choose $W_v \in \mathcal{W}_v$ such that the local L-factor equals the local zeta function at $g = 1$, i.e. such that

$$L(s, \pi_v \otimes \chi_v) = \int_{F_v^*} W_v \begin{pmatrix} x & 0 \\ 0 & 1 \end{pmatrix} \chi_v(x)|x|^{s-\frac{1}{2}} \, d^\times x \tag{37}$$

for any unramified quasi-character $\chi_v : F_v^* \to \mathbb{C}^*$ and $\mathrm{Re}(s) \gg 0$.

This is possible by [Ge75], Thm. 6.12 (ii); and by loc.cit., Prop. 6.17, W_v can be chosen such that SO(2) operates on W_v via ϱ_v for real archimedian v, and is "almost" $K_0(\mathfrak{c}(\pi_v))$-invariant for finite v.

For complex primes v of F, we can also choose a W_v satisfying (37) and which behaves well with respect to the SU(2)-action ϱ_v, as follows:

By [Kur77], there exists a function

$$\underline{W_v} = (W_v^0, W_v^1, W_v^2) : G(F_v) \to \mathbb{C}^3$$

such that $W_v^i \in \mathcal{W}_v$ for all i, and such that SU(2) operates by the right via ϱ_v on $\underline{W_v}$; i.e. for all $g \in G(F_v)$ and $h \in \mathrm{SU}(2)$, we have

$$\underline{W_v}(gh) = \underline{W_v}(g)\varrho_{\mathbb{C}}(h).$$

Note that W_v^1 is thus invariant under right multiplication by a diagonal matrix $\begin{pmatrix} u & 0 \\ 0 & \overline{u} \end{pmatrix}$ with $u \in S^1 \subseteq \mathbb{C}$. Since π_v has trivial central character for archimedian v by our assumption, a function in $\mathcal{W}_v$ is also invariant under $Z(F_v)$. Thus we have

$$W_v^1 \left(g \begin{pmatrix} u & 0 \\ 0 & 1 \end{pmatrix} \right) = W_v^1(g) \quad \text{for all } g \in G(F_v),\ u \in S^1.$$

W_v^1 can be described explicitly in terms of a certain Bessel function, as follows. The modified Bessel differential equation of order $\alpha \in \mathbb{C}$ is

$$x^2 \frac{d^2y}{dx^2} + x\frac{dy}{dx} - (x^2 + \alpha^2)y = 0.$$

Its solution space (on $\{\mathrm{Re}\, z > 0\}$) is two-dimensional; we are only interested in the second standard solution K_v, which is characterised by the asymptotics

$$K_v(z) \sim \sqrt{\frac{\pi}{2z}}\, e^{-z}$$

(cf. [We71]). By [Kur77],[3] we have $W_v^1 \begin{pmatrix} x & 0 \\ 0 & 1 \end{pmatrix} = \frac{2}{\pi} x^2 K_0(4\pi x)$.
(W_v^0 and W_v^2 can also be described in terms of Bessel functions; they are linearly dependent and scalar multiples of $x^2 K_1(4\pi x)$.)
By [JL70], Ch. 1, Thm. 6.2(vi), $\sigma(|\cdot|_{\mathbb{C}}^{1/2}, |\cdot|_{\mathbb{C}}^{-1/2}) \cong \pi(\mu_1, \mu_2)$ with

$$\mu_1(z) = z^{1/2}\overline{z}^{\,-1/2} = |z|_{\mathbb{C}}^{-1/2} z, \qquad \mu_2(z) = z^{-1/2}\overline{z}^{1/2} = |z|_{\mathbb{C}}^{-1/2}\overline{z};$$

and the L-series of the representation is the product of the L-factors of these two characters:

$$\begin{aligned} L_v(s, \pi_v) = L(s, \mu_1) L(s, \mu_2) &= 2\,(2\pi)^{-(s+\frac{1}{2})}\Gamma(s+\tfrac{1}{2}) \cdot 2\,(2\pi)^{-(s+\frac{1}{2})}\Gamma(s+\tfrac{1}{2}) \\ &= 4\,(2\pi)^{-(2s+1)}\Gamma(s+\tfrac{1}{2})^2. \end{aligned}$$

On the other hand, letting $d^\times x = \frac{dx}{|x|_{\mathbb{C}}} = \frac{dr}{r} d\vartheta$ (for $x = re^{i\vartheta}$), we have for $\mathrm{Re}(s) > -\frac{1}{2}$:

$$\begin{aligned} \int_{\mathbb{C}^*} W_v^1 \begin{pmatrix} x & 0 \\ 0 & 1 \end{pmatrix} |x|_{\mathbb{C}}^{s-\frac{1}{2}}\, d^\times x &= \int_{S^1}\int_{\mathbb{R}_+} W_v^1 \begin{pmatrix} re^{i\vartheta} & 0 \\ 0 & 1 \end{pmatrix} |x|_{\mathbb{C}}^{s-\frac{1}{2}}\, \frac{dr}{r}\, d\vartheta \\ &= 4\int_0^\infty x^2 K_0(4\pi x) x^{2s-1}\, \frac{dx}{x} \\ &\qquad \text{(invariance under } \mathrm{SU}(2)\cdot Z(F_v) \text{ gives a constant integral w.r.t. } \vartheta) \\ &= 4\,(4\pi)^{-2s+1} \int_0^\infty K_0(x) x^{2s}\, dx \\ &= 4\,(4\pi)^{-2s+1}\, 2^{2s-1}\, \Gamma(s+\tfrac{1}{2})^2 \\ &= 4\,(2\pi)^{-2s+1}\, \Gamma(s+\tfrac{1}{2})^2 \end{aligned}$$

by ([DLMF] 10.43.19). Thus we have

$$\int_{\mathbb{C}^*} W_v^1 \begin{pmatrix} x & 0 \\ 0 & 1 \end{pmatrix} |x|_{\mathbb{C}}^{s-\frac{1}{2}}\, d^\times x = (2\pi)^2\, L_v(s, \pi_v)$$

for all $\mathrm{Re}(s) > -\frac{1}{2}$.We set $W_v := (2\pi)^{-2}\, W_v^1$; thus (37) holds also for complex primes.
Now that we have defined W_v for all primes v, we put $W^p(g) := \prod_{v\nmid p} W_v(g_v)$ for all $g = (g_v)_v \in G(\mathbb{A}^p)$. We will also need the vector-valued function $\underline{W}^p : G(\mathbb{A}_F) \to V$ given by

$$\underline{W}^p(g) := \prod_{v\nmid p \text{ finite or } v \text{ real}} W_v(g_v) \cdot \bigotimes_{v \text{ complex}} (2\pi)^{-2}\underline{W_v}(g_v).$$

[3]Note that [Kur77] uses a slightly different definition of the K_v, which is $\frac{2}{\pi}$ times our K_v.

4.6 p-adic measures of automorphic forms

Now return to our $\pi \in \mathfrak{A}_0(G, \underline{2}, \chi_Z, \underline{\alpha_1}, \underline{\alpha_2})$. We fix an additive character $\psi : \mathbb{A} \to \mathbb{C}^*$ which is trivial on F, and let ψ_v denote the restriction of ψ to $F_v \hookrightarrow \mathbb{A}$, for all primes v. We further require that $\ker(\psi_\mathfrak{p}) \supseteq \mathcal{O}_\mathfrak{p}$ and $\mathfrak{p}^{-1} \not\subseteq \ker \psi_\mathfrak{p}$ for all $\mathfrak{p}|p$, so that we can apply the results of chapter 2.

As in chapter 2, let $\mu_{\pi_\mathfrak{p}} := \mu_{\alpha_{\mathfrak{p},1}/\nu_\mathfrak{p}} = \mu_{q_\mathfrak{p}/\alpha_{\mathfrak{p},2}}$ denote the distribution $\chi_{q_\mathfrak{p}/\alpha_{\mathfrak{p},2}}(x)\psi_\mathfrak{p}(x)dx$ on $F_\mathfrak{p}$, and let $\mu_{\pi_p} := \prod_{\mathfrak{p}|p} \mu_{\pi_\mathfrak{p}}$ be the product distribution on $F_p := \prod_{\mathfrak{p}|p} F_\mathfrak{p}$.

Define $\phi = \phi_\pi : \mathfrak{Co}(F_{S_1} \times F^*_{S_2}) \times \mathbb{I}^p \to \mathbb{C}$ by

$$\phi(U, x^p) := \sum_{\zeta \in F^*} \mu_{\pi_p}(\zeta U) W^p \begin{pmatrix} \zeta x^p & 0 \\ 0 & 1 \end{pmatrix}.$$

By proposition 2.13(a), we have for each $U \in \mathfrak{Co}(F_{S_1} \times F^*_{S_2})$:

$$\begin{aligned} \phi_U(x) := \phi(x_p U, x^p) &= \sum_{\zeta \in F^*} \mu_{\pi_p}(\zeta x_p U) W^p \begin{pmatrix} \zeta x^p & 0 \\ 0 & 1 \end{pmatrix} \\ &= \sum_{\zeta \in F^*} W \begin{pmatrix} \zeta x & 0 \\ 0 & 1 \end{pmatrix}, \end{aligned}$$

where $W(g) := W_U(g_p) W^p(g^p)$ lies in the global Whittaker model $\mathcal{W} = \mathcal{W}(\pi)$ for all $g = (g_p, g^p) \in G(\mathbb{A})$, putting $W_U := W_{1_U}$; so ϕ is well-defined and lies in $\mathcal{D}(S_1, \mathbb{C})$ (since W is smooth and rapidly decreasing; distribution property, F^*- and $U^{p,\infty}$-invariance being clear by the definitions of ϕ and W^p).

Let $\mu_\pi := \mu_{\phi_\pi}$ be the distribution on $\mathcal{G}_p$ corresponding to ϕ_π, as defined in (22), and let $\kappa_\pi := \kappa_{\phi_\pi} \in H^d(F^{*\prime}, \mathcal{D}_f(S_1, \mathbb{C}))$ be the cohomology class defined by (23) and (24).

Theorem 4.12. *Let $\pi \in \mathfrak{A}_0(G, \underline{2}, \chi_Z, \underline{\alpha_1}, \underline{\alpha_2})$; we assume the $\alpha_{\mathfrak{p},i}$ to be ordered such that $|\alpha_{\mathfrak{p},1}| \le |\alpha_{\mathfrak{p},2}|$ for all $\mathfrak{p}|p$. (So $\chi_{\mathfrak{p},1} = |\cdot|\chi_{\mathfrak{p},2}$ for all special $\pi_\mathfrak{p}$.)*
(a) Let $\chi : \mathcal{G}_p \to \mathbb{C}^*$ be a character of finite order with conductor $\mathfrak{f}(\chi)$. Then we have the interpolation property

$$\int_{\mathcal{G}_p} \chi(\gamma) \mu_\pi(d\gamma) = \tau(\chi) \prod_{\mathfrak{p} \in S_p} e(\pi_\mathfrak{p}, \chi_\mathfrak{p}) \cdot L(\tfrac{1}{2}, \pi \otimes \chi),$$

where

$$e(\pi_\mathfrak{p}, \chi_\mathfrak{p}) = \begin{cases} \dfrac{(1 - \alpha_{\mathfrak{p},1} x_\mathfrak{p} q_\mathfrak{p}^{-1})(1 - \alpha_{\mathfrak{p},2} x_\mathfrak{p}^{-1} q_\mathfrak{p}^{-1})(1 - \alpha_{\mathfrak{p},2} x_\mathfrak{p} q_\mathfrak{p}^{-1})}{(1 - x_\mathfrak{p} \alpha_{\mathfrak{p},2}^{-1})}, & \mathrm{ord}_\mathfrak{p}(\mathfrak{f}(\chi)) = 0 \\ & \textit{and } \pi \textit{ spherical}, \\ \dfrac{(1 - \alpha_{\mathfrak{p},1} x_\mathfrak{p} q_\mathfrak{p}^{-1})(1 - \alpha_{\mathfrak{p},2} x_\mathfrak{p}^{-1} q_\mathfrak{p}^{-1})}{(1 - x_\mathfrak{p} \alpha_{\mathfrak{p},2}^{-1})}, & \mathrm{ord}_\mathfrak{p}(\mathfrak{f}(\chi)) = 0 \\ & \textit{and } \pi \textit{ special}, \\ (\alpha_{\mathfrak{p},2}/q_\mathfrak{p})^{\mathrm{ord}_\mathfrak{p}(\mathfrak{f}(\chi))}, & \mathrm{ord}_\mathfrak{p}(\mathfrak{f}(\chi)) > 0 \end{cases}$$

and $x_\mathfrak{p} := \chi_\mathfrak{p}(\varpi_\mathfrak{p})$.
(b) κ_π *is integral (cf. definition 3.12). For* $\underline{\mu} \in \Sigma$, *let* $\kappa_{\pi,\underline{\mu}}$ *be the projection of* κ_π *to* $H^d(F^{*\prime}, \mathcal{D}_f(S_1, \mathbb{C}))_{\pi,\underline{\mu}}$. *Then* $\kappa_{\pi,\underline{\mu}}$ *is integral of rank* ≤ 1.

Proof. (a) We consider χ as a character on $\mathbb{I}_F/F^*$, and choose a subgroup $V = \prod_{\mathfrak{p}|p} V_\mathfrak{p} \subseteq U_p$ such that $\chi_p|_V = 1$.
Since π is unitary, we have $|\alpha_{\mathfrak{p},2}| \geq \sqrt{q_\mathfrak{p}} > 1 = |\chi_\mathfrak{p}(\varpi_\mathfrak{p})|$ for all $\mathfrak{p}$, thus $e(\pi_\mathfrak{p}, \chi_\mathfrak{p}|\cdot|_\mathfrak{p}^s)$ is non-singular for all $s \geq 0$, and we will be able to apply proposition 2.4 locally below.
We have

$$\int_{\mathcal{G}_p} \chi(\gamma)\mu_\pi(d\gamma) = [U_p : V] \int_{\mathbb{I}_F/F^*} \chi(x)\phi_V(x)d^\times x,$$

and therefore we have to show that the equality

$$[U_p : V] \int_{\mathbb{I}_F/F^*} \chi(x)|x|^s\phi_V(x)d^\times x = N(\mathfrak{f}(\chi))^s\tau(\chi)\prod_{\mathfrak{p}|p} e(\pi_\mathfrak{p}, \chi_\mathfrak{p}|\cdot|_\mathfrak{p}^s)\cdot L(s+\tfrac{1}{2}, \pi\otimes\chi)$$

holds for $s = 0$. Since both the left-hand side and $L(s+\frac{1}{2}, \pi\otimes\chi)$ are holomorphic in s (cf. [Ge75], Thm. 6.18), it suffices to show this for $\mathrm{Re}(s) \gg 0$. But for such s, we have

$$[U_p : V] \int_{\mathbb{I}_F/F^*} \chi(x)|x|^s\phi_V(x)d^\times x = \int_{\mathbb{I}_F} \chi(x)|x|^s W\begin{pmatrix} x & 0 \\ 0 & 1 \end{pmatrix} d^\times x$$

$$\begin{aligned}
&= [U_p : V] \int_{F_p^*} \chi_p(x)|x|^s W_V\begin{pmatrix} x & 0 \\ 0 & 1 \end{pmatrix} d^\times x \cdot \int_{\mathbb{I}_F^p} \chi^p(y)|y|^s W^p\begin{pmatrix} y & 0 \\ 0 & 1 \end{pmatrix} d^\times y \\
&= \prod_{\mathfrak{p}|p} \int_{F_\mathfrak{p}^*} \chi_\mathfrak{p}(x)|x|_\mathfrak{p}^s \mu_{\pi_\mathfrak{p}}(dx) \cdot L_{S_p}(s+\tfrac{1}{2}, \pi\otimes\chi) \\
&= \prod_{\mathfrak{p}|p} \left(e(\pi_\mathfrak{p}, \chi_\mathfrak{p}|\cdot|_\mathfrak{p}^s)\tau(\chi_\mathfrak{p}|\cdot|_\mathfrak{p}^s)\right) \cdot L(s+\tfrac{1}{2}, \pi\otimes\chi) \\
&= N(\mathfrak{f}(\chi))^s\tau(\chi)\prod_{\mathfrak{p}|p} e(\pi_\mathfrak{p}, \chi_\mathfrak{p}|\cdot|_\mathfrak{p}^s) \cdot L(s+\tfrac{1}{2}, \pi\otimes\chi)
\end{aligned}$$

by propositions 2.13, 2.4 and equation (37).
(b) Let $\lambda_{\underline{\alpha_1},\underline{\alpha_2}} \in \mathcal{B}^{\underline{\alpha_1},\underline{\alpha_2}}(F_p, \mathbb{C})$ be the image of $\otimes_{v|p}\lambda_{a_v,\nu_v}$ under the map (13). For each $\psi \in \mathcal{B}_{\underline{\alpha_1},\underline{\alpha_2}}(F_p, \mathbb{C})$, define

$$\begin{aligned}
\langle\Phi_\pi, \psi\rangle(g^p, g_p) &:= \sum_{\zeta\in F^*} \lambda_{\underline{\alpha_1},\underline{\alpha_2}}\left(\begin{pmatrix} \zeta & 0 \\ 0 & 1 \end{pmatrix} g_p \cdot \psi\right) \underline{W}^p\left(\begin{pmatrix} \zeta & 0 \\ 0 & 1 \end{pmatrix} g^p\right) \\
&=: \sum_{\zeta\in F^*} \underline{W_\psi}\left(\begin{pmatrix} \zeta & 0 \\ 0 & 1 \end{pmatrix} g\right)
\end{aligned}$$

for a V-valued function $\underline{W_\psi}$ whose every coordinate function is in $\mathcal{W}(\pi)$.
This defines a map $\Phi_\pi : G(\mathbb{A}^p) \to \mathcal{B}^{\underline{\alpha_1},\underline{\alpha_2}}(F_p, V)$. In fact, Φ_π lies in $S_2(G, \mathfrak{m}, \underline{\alpha_1}, \underline{\alpha_2})$, where $\mathfrak{m}$ is the prime-to-p part of $\mathfrak{f}(\pi)$:
Condition (a) of definition 4.5 follows from the fact that the W_v are almost $K_0(\mathfrak{c}(\pi_v))$-invariant, for $v \nmid p, \infty$. For condition (b), we check that $\langle \Phi_\pi, \psi\rangle$ satisfies the conditions (i)-(v) in the definition of $\mathcal{A}_0(G, \mathrm{harm}, \underline{2}, \chi)$:
Each coordinate function of $\langle \Phi_\pi, \psi\rangle$ lies in (the underlying space of) π by [Bu98], Thm. 3.5.5, thus $\langle \Phi, \psi\rangle$ fulfills (i) and (v), and has moderate growth. (ii) and (iv) follow from the choice of the W_v and $\underline{W_v}$. Now since $\pi_v \cong \sigma(|\cdot|_v^{1/2}, |\cdot|_v^{-1/2})$ for $v|\infty$, $\langle \Phi, \psi\rangle|_{B'_{F_v}} \cdot \underline{\beta_v} = C\sum_{\zeta\in F^*} \underline{W_v}\left(\begin{smallmatrix}\zeta t & 0\\ 0 & 1\end{smallmatrix}\right) \cdot \underline{\beta_v}$ is harmonic for each archimedean place v of F: for real v, it is well-known that $f(z)/y$ is holomorphic for $f \in \mathcal{D}(2)$, and thus $f \cdot (\beta_v)_1$ is harmonic; for complex v, harmonicity follows from the other conditions, see e.g. [Kur78], p. 546 or [We71].
An easy calculation shows that

$$\lambda_{\underline{\alpha_1},\underline{\alpha_2}}\left(\begin{pmatrix}\zeta & 0\\ 0 & 1\end{pmatrix}\delta_{\underline{\alpha_1},\underline{\alpha_2}}(1_U)\right) = \int_{\zeta U}\prod_{\mathfrak{p}|p}\chi_{\alpha_{\mathfrak{p}},2}(-x)\psi_{\mathfrak{p}}(-x)dx = \mu_{\pi_p}(\zeta U)$$

for all $\zeta \in F^*$, and therefore we have

$$\begin{aligned}\Delta^{\underline{\alpha_1},\underline{\alpha_2}}(\Phi_\pi)(U, x^p) &= \sum_{\zeta\in F^*}\lambda_{\underline{\alpha_1},\underline{\alpha_2}}\left(\begin{pmatrix}\zeta & 0\\ 0 & 1\end{pmatrix}\delta_{\underline{\alpha_1},\underline{\alpha_2}}(1_U)\right)W^p\begin{pmatrix}\zeta x^p & 0\\ 0 & 1\end{pmatrix}\\ &= \sum_{\zeta\in F^*}\mu_{\pi_p}(\zeta U)W^p\begin{pmatrix}\zeta x^p & 0\\ 0 & 1\end{pmatrix} = \phi_\pi(U, x^p).\end{aligned}$$

Let R be the integral closure of $\mathbb{Z}[a_\mathfrak{p}, \nu_\mathfrak{p}; \mathfrak{p}|p]$ in its field of fractions; thus R is a Dedekind ring $\subseteq \overline{\mathcal{O}}$ for which $\mathcal{B}_{\underline{\alpha_1},\underline{\alpha_2}}(F, R)$ is defined. Since $\mathbb{C}$ is a flat R-module,

$$H^d(G(F)^+, \mathcal{A}_f(\mathfrak{m}, \underline{\alpha_1}, \underline{\alpha_2}, \mathcal{M}, R)) \otimes \mathbb{C} \to H^d(G(F)^+, \mathcal{A}_f(\mathfrak{m}, \underline{\alpha_1}, \underline{\alpha_2}, \mathcal{M}, \mathbb{C}))$$

is an isomorphism by proposition 4.8. The map (36) can be described as the "R-valued" map

$$H^d(G(F)^+, \mathcal{A}_f(\mathfrak{m}, \underline{\alpha_1}, \underline{\alpha_2}, \mathcal{M}, R)) \to H^d\big(F^{*\prime}, \mathcal{D}_f(R)\big)$$

tensored with $\mathbb{C}$. By proposition 4.11, κ_π lies in its image, and thus in $H^d\big(F^{*\prime}, \mathcal{D}_f(R)\big) \otimes \mathbb{C}$; i.e. it is integral.
Similarly, it follows from propositions 4.8 and 4.10 that $\kappa_{\pi,\underline{\mu}}$ is integral of rank ≤ 1. ☐

COROLLARY 4.13. *μ_π is a p-adic measure.*

Proof. By proposition 3.8, $\mu_\pi = \mu_{\phi_\pi} = \mu_{\kappa_\pi}$. Since κ_π is integral, μ_{κ_π} is a p-adic measure by corollary 3.13. ☐

4.7 Vanishing order of the p-adic L-function

Let $L_1, \dots, L_t$ be independent $\mathbb{Z}_p$-extensions of F, and let $\ell_1, \dots, \ell_t : \mathcal{G}_p \to p^{\varepsilon_p}\mathbb{Z}_p$ be the homomorphisms corresponding to them (as in section 3.2). Then we have the *p-adic L-function*

$$L_p(\underline{s}, \pi) := L_p(\underline{s}, \kappa_\pi) := L_p(s_1, \dots, s_t, \kappa_{\pi,+}) := \int_{\mathcal{G}_p} \prod_{i=1}^{t} \exp_p(s_i \ell_i(\gamma)) \mu_\pi(d\gamma)$$

of definition 3.5, with $s_1, \dots, s_t \in \mathbb{Z}_p$. $L_p(\underline{s}, \pi)$ is a locally analytic function with values in the one-dimensional $\mathbb{C}_p$-vector space $V_{\kappa_{\pi,+}} = L_{\kappa,\overline{\mathcal{O}},+} \otimes_{\overline{\mathcal{O}}} \mathbb{C}_p$. By theorem 3.11, we have

Theorem 4.14. *$L_p(\underline{s}, \pi)$ is a locally analytic (t-variabled) function, and all partial derivatives of order $\leq n := \#(S_1)$ vanish; i.e. we have*

$$\operatorname{ord}_{\underline{s}=\underline{0}} L_p(\underline{s}, \pi) \geq n.$$

Now let E be a modular elliptic curve over F, corresponding to an automorphic representation π; by this we mean that the local L-factors of the Hasse-Weil L-function $L(E, s)$ and of the automorphic L-function $L(s - \frac{1}{2}, \pi)$ coincide at all places v of F. From the definition of the respective L-factors (cf. [Si86] for the Hasse-Weil L-function, [Ge75] for the automorphic L-function) we know that π has trivial central character. Moreover, for $\mathfrak{p}|p$, $\pi_\mathfrak{p}$ is a principal series representation iff E has good reduction at $\mathfrak{p}$, and in this case $\pi_\mathfrak{p}$ is ordinary iff E is ordinary (i.e. not supersingular) at $\mathfrak{p}$; $\pi_\mathfrak{p}$ is a special (resp. Steinberg) representation iff E has multiplicative (resp. split multiplicative) reduction at $\mathfrak{p}$. For $v|\infty$, π_v is "of weight 2" as assumed before.
We say that E is *p-ordinary* if it has good ordinary or multiplicative reduction at all places $\mathfrak{p}|p$ of F. So E is p-ordinary iff π is ordinary at all $\mathfrak{p}|p$. In this case, we define the p-adic L-function of E by $L_p(E, \underline{s}) := L_p(\underline{s}, \pi)$.
For each $i \in \{1, \dots, t\}$ and each prime $\mathfrak{p}|p$ of F, we write $\ell_{\mathfrak{p},i}$ for the restriction of ℓ_i to $F_\mathfrak{p} \hookrightarrow \mathbb{I} \twoheadrightarrow \mathcal{G}_p$. Let $q_\mathfrak{p}$ be the Tate period of $E|F_\mathfrak{p}$ and $\operatorname{ord}_\mathfrak{p}$ the normalized valuation on $F_\mathfrak{p}^*$. Defining the *L-invariants* of $E|F_\mathfrak{p}$ with respect to L_i by

$$\mathcal{L}_{\mathfrak{p},i}(E) := \ell_{\mathfrak{p},i}(q_\mathfrak{p}) / \operatorname{ord}_\mathfrak{p}(q_\mathfrak{p}),$$

we can generalize Hida's exceptional zero conjecture to general number fields:

Conjecture 4.15. *Let S_1 be the set of $\mathfrak{p}|p$ at which E has split multiplicative reduction, $n := \#S_1$, $S_2 := S_p \setminus S_1$. Then*

$$\operatorname{ord}_{\underline{s}=\underline{0}} L_p(E, \underline{s}) \geq n, \tag{38}$$

and we have

$$\frac{\partial^n}{\partial s_i^n} L_p(E, \underline{s})|_{\underline{s}=\underline{0}} = n! \prod_{\mathfrak{p} \in S_1} \mathcal{L}_{\mathfrak{p},i}(E) \prod_{\mathfrak{p} \in S_2} e(\pi_\mathfrak{p}, 1) \cdot L(E, 1), \tag{39}$$

for all $i = 1, \ldots, t$, where $e(\pi_\mathfrak{p}, 1) = (1 - \alpha_{\mathfrak{p},1}{}^{-1})^2$ if E has good ordinary reduction at $\mathfrak{p}$, and $e(\pi_\mathfrak{p}, 1) = 2$ if E has non-split multiplicative reduction at $\mathfrak{p}$.

Note that the conjecture (when considered for all sets of independent $\mathbb{Z}_p$-extensions of F) also determines the "mixed" partial derivatives $\frac{\partial^k}{\partial^{\underline{n}} \underline{s}} L_p(E, \underline{0})$ of order n, since they can be written as $\mathbb{Q}$-linear combinations of n-th "pure" partial derivatives $\frac{\partial^n}{\partial {s'}_i^n} L_p(E, \underline{0})$ with respect to other choices of independent $\mathbb{Z}_p$-extensions of F (cf. the proof of proposition 3.9).
Theorem 4.14 immediately implies the first part (38) of the conjecture:

Corollary 4.16. *Let E be a p-ordinary modular elliptic curve over F. Let n be the number of places $\mathfrak{p}|p$ at which E has split multiplicative reduction. Then we have*

$$\operatorname{ord}_{\underline{s}=\underline{0}} L_p(E, \underline{s}) \geq n.$$

In future work, we hope to also establish formula (39) for a class of non-totally-real number fields.

References

[BL95] L. Barthel, R. Livne: *Modular Representations of* GL_2 *of a Local Field: The Ordinary, Unramified Case.* Journal of Number Theory 55, 1-27 (1995).

[By98] J. Bygott: *Modular Forms and Modular Sumbols over Imaginary Quadratic Fields.* PhD thesis, 1998.

[Br82] K. Brown: *Cohomology of Groups.* Graduate Texts in Mathematics, Springer-Verlag, 1982.

[Bu98] D. Bump: *Automorphic Forms and Representations.* Cambridge University Press, 1998.

[CKM04] J. Cogdell, H. Kim, M. Murty: *Lectures on automorphic L-functions.* Fields Institute Monographs, 20. American Mathematical Society, Providence, RI, 2004.

[De13] H. Deppe: *p-adic L-functions of automorphic forms*, Dissertation (Ph.D. thesis) at Bielefeld University, 2013, `http://pub.uni-bielefeld.de/publication/2629389`.

[DLMF] NIST Digital Library of Mathematical Functions, `http://dlmf.nist.gov/`.

[Ge75] S. Gelbart: *Automorphic forms on adele groups.* Princeton University Press and University of Tokyo Press, 1975.

[GS93] R. Greenberg, G. Stevens: *p-adic L-functions and p-adic periods of modular forms.* Invent. Math. 111 (1993), 407-447.

[Har87] Sh. Haran: *p-adic L-functions for modular forms.* Compositio Math. 62 (1987), no. 1, 3146.

[Ha87] G. Harder: *Eisenstein cohomology of arithmetic groups. The case GL2.* Invent. Math. 89 (1987), 37-118.

[Hi09] H. Hida: *L-invariants of Tate curves.* Pure Appl. Math. Q. 5, 1343-1384 (2009)

[JL70] H. Jacquet; R. P. Langlands: *Automorphic forms on GL(2).* Lecture Notes in Mathematics, Vol. 114. Springer-Verlag, Berlin-New York, 1970.

[Kud04] St. Kudla: *From Modular Forms to Automorphic Representations.* In: Bump, Cogdell et al.: *An Introduction to the Langlands Program*, Birkhäuser Boston, 2004.

[Kur77] P. F. Kurcanov: *The zeta-function of elliptic curves over certain abelian extensions of imaginary quadratic fields.* English translation in: Math. USSR-Sb. 31 (1977), no. 1, 49-62 (1978).

[Kur78] P. F. Kurcanov: *Cohomology of discrete groups and Dirichlet series connected with Jacquet-Langlands cusp forms.* English translation in: Math. USSR Izvestija, Vol 12 (1978), No. 3, 543-555.

[MTT86] B. Mazur, J. Tate and J. Teitelbaum: *On p-adic analogues of the conjectures of Birch and Swinnerton-Dyer.* Invent. Math. 84 (1986), 1-48.

[Neu92] J. Neukirch: *Algebraische Zahlentheorie.* Springer-Lehrbuch, 1992.

[Se71] J.-P. Serre: *Cohomologie des groupes discrets.* Ann. of Math. Studies 70 (1971), 77-169.

[Se80] J.-P. Serre: *Trees.* Springer Monographs in Mathematics, 1980.

[Si86] J. Silverman: *Arithmetic of Elliptic Curves.* Graduate Texts in Mathematics, Springer-Verlag, 1986.

[Sp14] M. Spieß: *On special zeros of p-adic L-functions of Hilbert modular forms.* Invent. Math. (2014), 196: 69-138.

[Wa82] L.C. Washington: *Introduction to Cyclotomic Fields.* Graduate Texts in Mathematics, Springer-Verlag, 1982.

[We71] A. Weil: *Dirichlet Series and Automorphic Forms.* Lecture Note in Mathematics, Springer-Verlag, 1971.

Holger Deppe
Mathematical Institute
University of Cologne
Weyertal 86-90
50931 Cologne
Germany

Hyperplane Mass Partitions via Relative Equivariant Obstruction Theory

Pavle V. M. Blagojević, Florian Frick,
Albert Haase, Günter M. Ziegler[1]

Received: October 6, 2015
Revised: April 28, 2016

Communicated by Christian Bär

Abstract. The Grünbaum–Hadwiger–Ramos hyperplane mass partition problem was introduced by Grünbaum (1960) in a special case and in general form by Ramos (1996). It asks for the "admissible" triples (d, j, k) such that for any j masses in $\mathbb{R}^d$ there are k hyperplanes that cut each of the masses into 2^k equal parts. Ramos' conjecture is that the Avis–Ramos necessary lower bound condition $dk \geq j(2^k - 1)$ is also sufficient. We develop a "join scheme" for this problem, such that non-existence of an $\mathfrak{S}_k^{\pm}$-equivariant map between spheres $(S^d)^{*k} \to S(W_k \oplus U_k^{\oplus j})$ that extends a test map on the subspace of $(S^d)^{*k}$ where the hyperoctahedral group $\mathfrak{S}_k^{\pm}$ acts non-freely, implies that (d, j, k) is admissible. For the sphere $(S^d)^{*k}$ we obtain a very efficient regular cell decomposition, whose cells get a combinatorial interpretation with respect to measures on a modified moment curve. This allows us to apply relative equivariant obstruction theory successfully, even in the case when the difference of dimensions of the spheres $(S^d)^{*k}$ and $S(W_k \oplus U_k^{\oplus j})$ is greater than one. The evaluation of obstruction classes leads to counting problems for concatenated Gray codes. Thus we give a rigorous, unified treatment of the previously announced cases of the Grünbaum–Hadwiger–Ramos problem, as well as a number of new cases for Ramos' conjecture.

2010 Mathematics Subject Classification: 55N25, 51N20, 52A35, 55R20

Keywords and Phrases: Hyperplane mass partition problem, equivariant topological combinatorics, equivariant obstruction theory

[1]P.B. was supported by the DFG via the Collaborative Research Center TRR 109 "Discretization in Geometry and Dynamics", and by the grant ON 174008 of Serbian Ministry of Education and Science.
F.F. and A.H. were supported by DFG via the Berlin Mathematical School.
G.M.Z. received funding from the ERC project no. 247029-*SDModels* and by DFG via the Research Training Group "Methods for Discrete Structures" and the Collaborative Research Center TRR 109 "Discretization in Geometry and Dynamics"

1 Introduction

1.1 The Grünbaum–Hadwiger–Ramos hyperplane mass partition problem

In 1960, Grünbaum [10, Sec. 4.(v)] asked whether for any convex body in $\mathbb{R}^k$ there are k affine hyperplanes that divide it into 2^k parts of equal volume: This is now known to be true for $k \leq 3$, due to Hadwiger [11] in 1966, and remains open and challenging for $k = 4$. (A weak partition result for $k = 4$ was given in 2009 by Dimitrijević-Blagojević [8].) For $k > 4$ it is false, as shown by Avis [1] in 1984 by considering a measure on a moment curve. In 1996, Ramos [15] proposed the following generalization of Grünbaum's problem.

The Grünbaum–Hadwiger–Ramos problem. *Determine the minimal dimension $d = \Delta(j, k)$ such that for every collection of j masses $\mathcal{M}$ on $\mathbb{R}^d$ there exists an arrangement of k affine hyperplanes $\mathcal{H}$ in $\mathbb{R}^d$ that equiparts $\mathcal{M}$.*

The Ham Sandwich theorem, conjectured by Steinhaus and proved by Banach, states that $\Delta(d, 1) = d$. The Grünbaum–Hadwiger–Ramos hyperplane mass partition problem was studied by many authors. It has been an excellent testing ground for different equivariant topology methods; see to our recent survey in [3].

The first general result about the function $\Delta(j, k)$ was obtained by Ramos [15], by generalizing Avis' observation: The lower bound

$$\Delta(j, k) \geq \tfrac{2^k-1}{k} j$$

follows from considering k measures with disjoint connected supports concentrated along a moment curve in $\mathbb{R}^d$. Ramos also conjectured that this lower bound is tight.

The Ramos conjecture. $\Delta(j, k) = \lceil \frac{2^k-1}{k} j \rceil$ *for every $j \geq 1$ and $k \geq 1$.*

All available evidence up to now supports this, though it has been established rigorously only in special cases.

1.2 Product scheme and join scheme

It seems natural to use $Y_{d,k} := (S^d)^k$ as a configuration space for any k oriented affine hyperplanes/halfspaces in $\mathbb{R}^d$, which leads to the following *product scheme*: If there is no equivariant map

$$(S^d)^k \longrightarrow_{\mathfrak{S}_k^{\pm}} S(U_k^{\oplus j})$$

from the configuration space to the unit sphere in the space $U_k^{\oplus j}$ of values on the orthants of $\mathbb{R}^k$ that sum to 0, which is equivariant with respect to the hyperoctahedral (signed permutation) group $\mathfrak{S}_k^{\pm}$, then there is no counterexample for the given parameters, so $\Delta(j, k) \leq d$.

However, our critical review [3] of the main papers on the Grünbaum–Hadwiger–Ramos problem since 1998 has shown that this scheme is very hard to handle: Except for the 2006 upper bounds by Mani-Levitska, Vrećica & Živaljević [13], derived from a Fadell–Husseini index calculation, it has produced very few valid results: The group action on $(S^d)^k$ is not free, the Fadell–Husseini index is rather large and thus yields weak results, and there is no efficient cell complex model at hand.

In this paper, we provide a new approach, which proves to be remarkably clean and efficient. For this, we use a *join scheme*, as introduced by Blagojević and Ziegler [4], which takes the form

$$F: \quad (S^d)^{*k} \longrightarrow_{\mathfrak{S}_k^\pm} S(W_k \oplus U_k^{\oplus j}).$$

Here the domain $(S^d)^{*k} \subseteq \mathbb{R}^{(d+1)\times k}$ is a sphere of dimension $dk + k - 1$, given by

$$X_{d,k} := \{(\lambda_1 x_1, \dots, \lambda_k x_k) : x_1, \dots, x_k \in S^d,\ \lambda_1, \dots, \lambda_k \geq 0,\ \lambda_1 + \cdots + \lambda_k = 1\},$$

where we write $\lambda_1 x_1 + \cdots + \lambda_k x_k$ as a short-hand for $(\lambda_1 x_1, \dots, \lambda_k x_k)$. The co-domain is a sphere of dimension $j(2^k - 1) + k - 2$. Both domain and co-domain are equipped with canonical $\mathfrak{S}_k^\pm$-actions. We observe that the map restricted to the points with non-trivial stabilizer (the "non-free part")

$$F': \quad X_{d,k}^{>1} \subset (S^d)^{*k} \longrightarrow_{\mathfrak{S}_k^\pm} S(W_k \oplus U_k^{\oplus j})$$

is the same up to homotopy for all test maps. If for any parameters (j, k, d) an equivariant extension F of F' does not exist, we get that $\Delta(j, k) \leq d$.

To decide the existence of this map, or at least obtain necessary criteria, we employ relative equivariant obstruction theory, as explained by tom Dieck [7, Sect. II.3]. This turns out to work beautifully, and have a few remarkable aspects:

- The Fox–Neuwirth [9]/Björner–Ziegler [2] combinatorial stratification method yields a simple and efficient cone stratification for the space $\mathbb{R}^{(d+1)\times k}$, which is equivariant with respect to the action of $\mathfrak{S}_k^\pm$ on the columns, and which respects the arrangement of k^2 subspaces of codimension d given by columns of a matrix $(x_1, \dots, x_d)$ being equal, opposite, or zero.
- This yields a small equivariant regular CW complex model for the sphere $(S^d)^{*k} \subseteq \mathbb{R}^{(d+1)\times k}$, for which the the non-free part, given by an arrangement of k^2 subspheres of codimension $d + 1$, is an invariant subcomplex. The cells $D_I^S(\sigma)$ in the complex are given by combinatorial data.
- To evaluate the obstruction cocycle, we use measures on a non-standard (binomial coefficient) moment curve. For the resulting test map, the relevant cells $D_I^S(\sigma)$ can be interpreted as k-tuples of hyperplanes such that some of the hyperplanes have to pass through prescribed points of the moment curve, or equivalently, they have to bisect some extra masses.

1.3 Statement of the main results

The join scheme reduces the Grünbaum–Hadwiger–Ramos problem to a combinatorial counting problem that can be solved by hand or by means of a computer: A k-bit *Gray code* is a $k \times 2^k$ binary matrix of all column vectors of length k such that two consecutive vectors differ by only one bit. Such a k-bit code can be interpreted as a Hamiltonian path in the graph of the k-cube $[0,1]^k$. The *transition count* of a row in a binary matrix A is the number of bit-changes, *not* counting a bit change from the last to the first entry. By *transition counts* of a matrix A we refer to the vector of the transition counts of the rows of the matrix A. Two binary matrices A and A' are *equivalent*, if A can be obtained from A' by a sequence of permutations of rows and/or inversion of bits in rows.

Definition 1.1. Let $d \geq 1$, $j \geq 1$, $\ell \geq 0$ and $k \geq 1$ be integers such that $dk = (2^k - 1)j + \ell$ with $0 \leq \ell \leq d-1$. A binary matrix A of size $k \times j2^k$ is an *ℓ-equiparting matrix* if

(a) $A = (A_1, \dots, A_j)$ for Gray codes $A_1, \dots, A_j$ with the property that the last column of A_i is equal to the first column of A_{i+1} for $1 \leq i < j$; and

(b) there is one row of the matrix A with the transition count $d - \ell$, while all other rows have transition count d.

Example 1.2. If $d = 5$, $j = 2$, $\ell = 1$ and $k = 3$, then a possible 1-equiparting matrix is

$$A = (A_1, A_2) = \begin{pmatrix} 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & \quad 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & \quad 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & \quad 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

In this example the first row of A has transition count 4 while the remaining two rows have transition count 5.

Theorem 1.3. *Let $d \geq 1$, $j \geq 1$, $\ell \geq 0$ and $k \geq 2$ be integers with the property that $dk = (2^k - 1)j + \ell$ and $0 \leq \ell \leq d-1$. The number of non-equivalent ℓ-equiparting matrices is the number of arrangements of k affine hyperplanes $\mathcal{H}$ that equipart a given collection of j disjoint intervals on a moment curve γ in $\mathbb{R}^d$, up to renumbering and orientation change of hyperplanes in $\mathcal{H}$, when it is forced that one of the hyperplanes passes through ℓ prescribed points on γ that lie to the left of the j disjoint intervals.*

In some situations this yields a solution for the Grünbaum–Hadwiger–Ramos problem.

Theorem 1.4. *Let $j \geq 1$ and $k \geq 3$ be integers, with $d := \lceil \frac{2^k-1}{k} j \rceil$ and $\ell := \lceil \frac{2^k-1}{k} j \rceil k - (2^k-1)j = dk - (2^k-1)j$, which implies $0 \leq \ell < k \leq d$. If the number of non-equivalent ℓ-equiparting matrices of size $k \times j2^k$ is odd, then*

$$\Delta(j,k) = \lceil \tfrac{2^k-1}{k} j \rceil.$$

Theorem 1.4 is also true for $k = 1$ (and thus $d = j$, $\ell = 0$), where it yields the Ham Sandwich theorem: In this case an equiparting matrix A is a row vector of length $2d$ and transition count d. Thus, each A_i is either $(0, 1)$ or $(1, 0)$, where A_i uniquely determines A_{i+1}. Hence, up to inversion of bits A is unique and so $\Delta(d, 1) \leq d$, and consequently $\Delta(d, 1) = d$.
While the situation for $k = 1$ hyperplane is fully understood, we seem to be far from a complete solution for the case of $k = 2$ hyperplanes. However, we do obtain the following instances.

Theorem 1.5. *Let $t \geq 1$. Then:*
(i) $\Delta(2^t - 1, 2) = 3 \cdot 2^{t-1} - 1$,
(ii) $\Delta(2^t, 2) = 3 \cdot 2^{t-1}$,
(iii) $\Delta(2^t + 1, 2) = 3 \cdot 2^{t-1} + 2$.

The statements (i) and (iii) were already known: Part (i) is the only case where the lower bound of Ramos and the upper bound of Mani-Levitska, Vrećica, and Živaljević [13, Thm. 39] coincide. Part (ii) is Hadwiger's result [11] for $t = 1$; the general case was previously claimed by Mani-Levitska et al. [13, Prop. 25]. However, the proof of the result was incorrect and not recoverable, as explained in [3, Sec. 8.1]. Here we recover this result by a different method of proof. Similarly, statement (iii) was claimed by Živaljević [17, Thm. 2.1] with a flawed proof; for an explanation of the gap see [3, Sec. 8.2], where we also gave a proof of (iii) via degrees of equivariant maps [3, Sec. 5]. Here we will prove all three cases of Theorem 1.5 in a uniform way.
In the case of $k = 3$ hyperplanes we prove using Theorem 1.4 the following instances of the Ramos conjecture.

Theorem 1.6.
(i) $\Delta(2, 3) = 5$,
(ii) $\Delta(4, 3) = 10$.

Statement (i) was previously claimed by Ramos [15, Sec. 6.1]. A gap in the method that Ramos developed and used to get this result was explained in [3, Sec. 7]. It is also claimed by Vrećica and Živaljević in the recent preprint [16] without a proof for the crucial [16, Prop. 3].
The reduction result of Hadwiger and Ramos $\Delta(j, k) \leq \Delta(2j, k - 1)$ applied to Theorem 1.6 implies the following consequences. For details on reduction results see for example [3, Sec. 3.3].

Corollary 1.7.
(i) $4 \leq \Delta(1, 4) \leq 5$,
(ii) $8 \leq \Delta(2, 4) \leq 10$.

Note that $\Delta(1, 4)$ is the open case for Grünbaum's original conjecture.

Acknowledgement

We are grateful to the referee for very careful and helpful comments on this paper.

2 The join configuration space test map scheme and equivariant obstruction theory

In this section we develop the join configuration test map scheme that was introduced in [5, Sec. 2.1]. A sufficient condition for $\Delta(j,k) \leq d$ will be phrased in terms of the non-existence of a particular equivariant map between representation spheres.

2.1 Arrangements of k hyperplanes

Let $\hat{H} = \{x \in \mathbb{R}^d : \langle x, v\rangle = a\}$ be an affine hyperplane determined by a vector $v \in \mathbb{R}^d \backslash \{0\}$ and a constant $a \in \mathbb{R}$. The hyperplane $\hat{H}$ determines two (closed) halfspaces

$$\hat{H}^0 = \{x \in \mathbb{R}^d : \langle x, v\rangle \geq a\} \qquad \text{and} \qquad \hat{H}^1 = \{x \in \mathbb{R}^d : \langle x, v\rangle \leq a\}.$$

Let $\mathcal{H} = (\hat{H}_1, \dots, \hat{H}_k)$ be an arrangement of k affine hyperplanes in $\mathbb{R}^d$, and let $\alpha = (\alpha_1, \dots, \alpha_k) \in (\mathbb{Z}/2)^k$. The *orthant* determined by the arrangement $\mathcal{H}$ and $\alpha \in (\mathbb{Z}/2)^k$ is the intersection

$$\mathcal{O}^{\mathcal{H}}_\alpha = \hat{H}_1^{\alpha_1} \cap \dots \cap \hat{H}_k^{\alpha_k}.$$

Let $\mathcal{M} = (\mu_1, \dots, \mu_j)$ be a collection of finite Borel probability measures on $\mathbb{R}^d$ such that the measure of each hyperplane is zero. Such measures will be called *masses.* The assumptions about the measures guarantee that $\mu_i(\hat{H}^0_s)$ depends continuously on $\hat{H}^0_s$.

An arrangement of affine hyperplanes $\mathcal{H} = (\hat{H}_1, \dots, \hat{H}_k)$ *equiparts* the collection of masses $\mathcal{M} = (\mu_1, \dots, \mu_j)$ if for every element $\alpha \in (\mathbb{Z}/2)^k$ and every $\ell \in \{1, \dots, j\}$

$$\mu_\ell(\mathcal{O}^{\mathcal{H}}_\alpha) = \tfrac{1}{2^k}.$$

2.2 The configuration spaces

The space of all oriented affine hyperplanes (or closed affine halfspaces) in $\mathbb{R}^d$ can be parametrized by the sphere S^d, where the north pole e_{d+1} and the south pole $-e_{d+1}$ represent hyperplanes at infinity. An oriented affine hyperplane in $\mathbb{R}^d$ *at infinity* is the set $\mathbb{R}^d$ or $\emptyset$, depending on the orientation. Indeed, embed $\mathbb{R}^d$ into $\mathbb{R}^{d+1}$ via the map $(\xi_1, \dots, \xi_d)^t \longmapsto (1, \xi_1, \dots, \xi_d)^t$. Then an oriented affine hyperplane $\hat{H}$ in $\mathbb{R}^d$ defines an oriented affine $(d-1)$-dimensional subspace of $\mathbb{R}^{d+1}$ that extends (uniquely) to an oriented linear hyperplane H in $\mathbb{R}^{d+1}$. The outer unit normal vector that determines the oriented linear hyperplane is a point on the sphere S^d.

We consider the following configuration spaces that parametrize arrangements of k oriented affine hyperplanes in $\mathbb{R}^d$:

(1) *The join configuration space*: $X_{d,k} := (S^d)^{*k} \cong S(\mathbb{R}^{(d+1)\times k})$,

(2) *The product configuration space*: $Y_{d,k} := (S^d)^k$.

The elements of the join $X_{d,k}$ can be presented as formal convex combinations $\lambda_1 v_1 + \dots + \lambda_k v_k$, where $\lambda_i \geq 0$, $\sum \lambda_i = 1$ and $v_i \in S^d$.

2.3 The group actions

The space of all ordered k-tuples of oriented affine hyperplanes in $\mathbb{R}^d$ has natural symmetries: Each hyperplane can change orientation and the hyperplanes can be permuted. Thus, the group $\mathfrak{S}_k^\pm := (\mathbb{Z}/2)^k \rtimes \mathfrak{S}_k$ encodes the symmetries of both configuration spaces.
The group $\mathfrak{S}_k^\pm$ acts on $X_{d,k}$ as follows. Each copy of $\mathbb{Z}/2$ acts antipodally on the appropriate sphere S^d in the join while the symmetric group $\mathfrak{S}_k$ acts by permuting factors in the join product. More precisely, for $((\beta_1,\ldots,\beta_k)\rtimes\pi)\in\mathfrak{S}_k^\pm$ and $\lambda_1 v_1+\cdots+\lambda_k v_k\in X_{d,k}$ the action is given by

$$((\beta_1,\ldots,\beta_k)\rtimes\tau)\cdot(\lambda_1 v_1+\cdots+\lambda_k v_k) = \\ \lambda_{\tau^{-1}(1)}(-1)^{\beta_1}v_{\tau^{-1}(1)}+\cdots+\lambda_{\tau^{-1}(k)}(-1)^{\beta_k}v_{\tau^{-1}(k)}.$$

The product space $Y_{d,k}$ is a subspace of the join $X_{d,k}$ via the diagonal embedding $Y_{d,k}\longrightarrow X_{d,k}, (v_1,\ldots,v_k)\longmapsto \frac{1}{k}v_1+\cdots+\frac{1}{k}v_k$. The product $Y_{d,k}$ is an invariant subspace of $X_{d,k}$ with respect to the $\mathfrak{S}_k^\pm$-action and consequently inherits the $\mathfrak{S}_k^\pm$-action from $X_{d,k}$. For $k\geq 2$, the action of $\mathfrak{S}_k^\pm$ is *not* free on either $X_{d,k}$ or $Y_{d,k}$.
The sets of points in the configuration spaces $X_{d,k}$ and $Y_{d,k}$ that have non-trivial stabilizer with respect to the action of $\mathfrak{S}_k^\pm$ can be described as follows:

$$X_{d,k}^{>1} = \{\lambda_1 v_1+\cdots+\lambda_k v_k : \\ \lambda_1\cdots\lambda_k=0,\ \text{or } \lambda_s=\lambda_r \text{ and } v_s=\pm v_r \text{ for some } 1\leq s<r\leq k\},$$

and

$$Y_{d,k}^{>1} = \{(v_1,\ldots,v_k) : v_s=\pm v_r \text{ for some } 1\leq s<r\leq k\}.$$

2.4 Test spaces

Consider the vector space $\mathbb{R}^{(\mathbb{Z}/2)^k}$, where the group element $((\beta_1,\ldots,\beta_k)\rtimes\tau)\in\mathfrak{S}_k^\pm$ acts on a vector $(y_{(\alpha_1,\ldots,\alpha_k)})_{(\alpha_1,\ldots,\alpha_k)\in(\mathbb{Z}/2)^k}\in\mathbb{R}^{(\mathbb{Z}/2)^k}$ by acting on its indices as

$$((\beta_1,\ldots,\beta_k)\rtimes\tau)\cdot(\alpha_1,\ldots,\alpha_k) = (\beta_1+\alpha_{\tau^{-1}(1)},\ldots,\beta_k+\alpha_{\tau^{-1}(k)}). \quad (1)$$

The subspace of $\mathbb{R}^{(\mathbb{Z}/2)^k}$ defined by

$$U_k = \Big\{(y_\alpha)_{\alpha\in(\mathbb{Z}/2)^k}\in\mathbb{R}^{(\mathbb{Z}/2)^k} : \sum_{\alpha\in(\mathbb{Z}/2)^k} y_\alpha = 0\Big\}$$

is $\mathfrak{S}_k^\pm$-invariant and therefore an $\mathfrak{S}_k^\pm$-subrepresentation.
Next we consider the vector space $\mathbb{R}^k$ and its subspace

$$W_k = \Big\{(z_1,\ldots,z_k)\in\mathbb{R}^k : \sum_{i=1}^k z_i = 0\Big\}.$$

The group $\mathfrak{S}_k^{\pm}$ acts on $\mathbb{R}^k$ by permuting coordinates, i.e., for $((\beta_1,\dots,\beta_k)\rtimes\tau)\in\mathfrak{S}_k^{\pm}$ and $(z_1,\dots,z_k)\in\mathbb{R}^k$ we have

$$((\beta_1,\dots,\beta_k)\rtimes\tau)\cdot(z_1,\dots,z_k)=(z_{\tau^{-1}(1)},\dots,z_{\tau^{-1}(k)}). \tag{2}$$

In particular, the subgroup $(\mathbb{Z}/2)^k$ of $\mathfrak{S}_k^{\pm}$ acts trivially on $\mathbb{R}^k$. The subspace $W_k\subset\mathbb{R}^k$ is $\mathfrak{S}_k^{\pm}$-invariant and consequently a $\mathfrak{S}_k^{\pm}$-subrepresentation.

2.5 Test maps

The product test map associated to the collection of j masses $\mathcal{M}=(\mu_1,\dots,\mu_j)$ from the configuration space $Y_{d,k}$ to the *test space* $U_k^{\oplus j}$ is defined by

$$\begin{aligned}\phi_{\mathcal{M}}\colon\qquad Y_{d,k}&\longrightarrow U_k^{\oplus j},\\ (v_1,\dots,v_k)&\longmapsto\Big(\big(\mu_i(H_{v_1}^{\alpha_1}\cap\dots\cap H_{v_k}^{\alpha_k})-\tfrac{1}{2^k}\big)_{(\alpha_1,\dots,\alpha_k)\in(\mathbb{Z}/2)^k}\Big)_{i\in\{1,\dots,j\}}.\end{aligned}$$

In this paper we mostly work with the join configuration space $X_{d,k}$. The corresponding join test map associated to a collection of j masses $\mathcal{M}=(\mu_1,\dots,\mu_j)$ maps the configuration space $X_{d,k}$ into the related *test space* $W_k\oplus U_k^{\oplus j}$. It is defined by

$$\begin{aligned}\psi_{\mathcal{M}}\colon\qquad X_{d,k}&\longrightarrow W_k\oplus U_k^{\oplus j},\\ \lambda_1v_1+\dots+\lambda_kv_k&\longmapsto(\lambda_1-\tfrac{1}{k},\dots,\lambda_k-\tfrac{1}{k})\oplus(\lambda_1\cdots\lambda_k)\cdot\phi_{\mathcal{M}}(v_1,\dots,v_k).\end{aligned}$$

Both maps $\phi_{\mathcal{M}}$ and $\psi_{\mathcal{M}}$ are $\mathfrak{S}_k^{\pm}$-equivariant with respect to the actions defined in Sections 2.3 and 2.4. Let $S(U_k^{\oplus j})$ and $S(W_k\oplus U_k^{\oplus j})$ denote the unit spheres in the vector spaces $U_k^{\oplus j}$ and $W_k\oplus U_k^{\oplus j}$, respectively. The maps $\phi_{\mathcal{M}}$ and $\psi_{\mathcal{M}}$ are called test maps since we have the following criterion, which reduces finding an equipartition to finding zeros of the test map.

Proposition 2.1. *Let $d\ge 1$, $k\ge 1$, and $j\ge 1$ be integers.*

(i) *Let $\mathcal{M}$ be a collection of j masses on $\mathbb{R}^d$, and let*

$$\phi_{\mathcal{M}}\colon Y_{d,k}\longrightarrow U_k^{\oplus j}\quad\text{and}\quad\psi_{\mathcal{M}}\colon X_{d,k}\longrightarrow W_k\oplus U_k^{\oplus j}$$

be the $\mathfrak{S}_k^{\pm}$-equivariant maps defined above. If $0\in\operatorname{im}\phi_{\mathcal{M}}$, or $0\in\operatorname{im}\psi_{\mathcal{M}}$, then there is an arrangement of k affine hyperplanes that equiparts $\mathcal{M}$.

(ii) *If there is no $\mathfrak{S}_k^{\pm}$-equivariant map of either type*

$$Y_{d,k}\longrightarrow S(U_k^{\oplus j})\quad\text{or}\quad X_{d,k}\longrightarrow S(W_k\oplus U_k^{\oplus j}),$$

then $\Delta(j,k)\le d$.

It is worth pointing out that $0 \in \operatorname{im} \phi_{\mathcal{M}}$ if and only if $0 \in \operatorname{im} \psi_{\mathcal{M}}$, while the existence of an $\mathfrak{S}_k^{\pm}$-equivariant map $Y_{d,k} \longrightarrow S(U_k^{\oplus j})$ implies the existence of a $\mathfrak{S}_k^{\pm}$-equivariant map $X_{d,k} \longrightarrow S(W_k \oplus U_k^{\oplus j})$ but not vice versa.
The homotopy class of the restrictions of the test maps $\phi_{\mathcal{M}}$ and $\psi_{\mathcal{M}}$ on the set of points with non-trivial stabilizer (as maps avoiding the origin) is independent of the choice of the masses $\mathcal{M}$, by the following proposition.

Proposition 2.2. *Let $\mathcal{M}$ and $\mathcal{M}'$ be collections of j masses in $\mathbb{R}^d$. Then*

(i) $0 \notin \operatorname{im} \phi_{\mathcal{M}}|_{Y_{d,k}^{>1}}$ *and* $0 \notin \operatorname{im} \psi_{\mathcal{M}}|_{X_{d,k}^{>1}}$,

(ii) $\phi_{\mathcal{M}}|_{Y_{d,k}^{>1}}$ *and* $\phi_{\mathcal{M}'}|_{Y_{d,k}^{>1}}$ *are* $\mathfrak{S}_k^{\pm}$*-homotopic as maps* $Y_{d,k}^{>1} \longrightarrow U_k^{\oplus j}\backslash\{0\}$, *and*

(iii) $\psi_{\mathcal{M}}|_{X_{d,k}^{>1}}$ *and* $\psi_{\mathcal{M}'}|_{X_{d,k}^{>1}}$ *are* $\mathfrak{S}_k^{\pm}$*-homotopic as maps* $X_{d,k}^{>1} \longrightarrow (W_k \oplus U_k^{\oplus j})\backslash\{0\}$.

Proof. (i) If $(v_1, \dots, v_k) \in Y_{d,k}^{>1}$, then $v_s = \pm v_r$ for some $1 \le s < r \le k$. Consequently, the corresponding hyperplanes H_{v_i} and H_{v_j} coincide, possibly with opposite orientations. Thus some of the orthants associated to the collection of hyperplanes $(H_{v_1}, \dots, H_{v_k})$ are empty. Consequently, Proposition 2.1 implies that $0 \notin \operatorname{im} \phi_{\mathcal{M}}|_{Y_{d,k}^{>1}}$.
In the case where $\lambda_1 v_1 + \dots + \lambda_k v_k \in X_{d,k}^{>1}$ the additional case $\lambda_s = 0$ for some $1 \le s \le k$ may occur. If $\lambda_s = 0$, then the s-th coordinate of $\psi(\lambda_1 v_1 + \dots + \lambda_k v_k) \in W_k \oplus U_k^{\oplus j}$ is equal to $-\frac{1}{k}$, and hence $0 \notin \operatorname{im} \psi_{\mathcal{M}}|_{X_{d,k}^{>1}}$.
(ii) The equivariant homotopy between $\phi_{\mathcal{M}}|_{Y_{d,k}^{>1}}$ and $\phi_{\mathcal{M}'}|_{Y_{d,k}^{>1}}$ is just the linear homotopy in $U_k^{\oplus j}$. For this the linear homotopy should not have zeros, compare [3, proof of Cor. 5.4]. It suffices to prove that for each point $(v_1, \dots, v_k) \in Y_{d,k}^{>1}$, the points $\phi_{\mathcal{M}}(v_1, \dots, v_k)$ and $\phi_{\mathcal{M}'}(v_1, \dots, v_k)$ belong to some affine subspace of the test space that is not linear.
First, observe that $\mathbb{R}^{(\mathbb{Z}/2)^k}$, considered as a real $(\mathbb{Z}/2)^k$ representation, is the real regular representation of $(\mathbb{Z}/2)^k$ and therefore it decomposes into the direct sum of all real irreducible representations. For this we use the fact that all real irreducible representations of $(\mathbb{Z}/2)^k$ are 1-dimensional. The subspace U_k seen as a real $(\mathbb{Z}/2)^k$ subrepresentation of $(\mathbb{Z}/2)^k$ decomposes as follows:

$$U_k \cong \bigoplus_{\alpha \in (\mathbb{Z}/2)^k \backslash\{0\}} V_\alpha. \tag{3}$$

Here V_α is the 1-dimensional real representation of $(\mathbb{Z}/2)^k$ determined by $\beta \cdot v = -v$ for $x \in V_\alpha$ if and only if $\alpha \cdot \beta := \sum \alpha_s \beta_s = 1 \in \mathbb{Z}/2$, for $\beta \in (\mathbb{Z}/2)^k$. The isomorphism (3) is given by the direct sum of the projections $\pi_\alpha \colon U_k \longrightarrow V_\alpha$, $\alpha \in (\mathbb{Z}/2)^k \backslash\{0\}$,

$$(y_\beta)_{\beta \in (\mathbb{Z}/2)^k \backslash\{0\}} \longmapsto \sum_{\alpha\cdot\beta=1} y_\beta - \sum_{\alpha\cdot\beta=0} y_\beta.$$

Now let $v_s = \pm v_r$. Consider $\alpha \in (\mathbb{Z}/2)^k$ given by $\alpha_s = 1 = \alpha_r$ and $\alpha_\ell = 0$ for $\ell \notin \{s, r\}$, and the corresponding projection $\pi_\alpha^{\oplus j}\colon U_k^{\oplus j} \longrightarrow V_\alpha^{\oplus j}$. Then

$$\pi_\alpha^{\oplus j} \circ \phi_{\mathcal{M}}(v_1, \dots, v_k) = \pi_\alpha^{\oplus j} \circ \phi_{\mathcal{M}'}(v_1, \dots, v_k) = (\pm 1, \dots, \pm 1).$$

(iii) Likewise, the linear homotopy between $\psi_{\mathcal{M}}|_{X_{d,k}^{>1}}$ and $\psi_{\mathcal{M}'}|_{X_{d,k}^{>1}}$ is equivariant and avoids zero. Let $\lambda_1 v_1 + \cdots + \lambda_k v_k \in X_{d,k}^{>1}$. If $\lambda := \lambda_1 \cdots \lambda_k \neq 0$, $\lambda_s = \lambda_r$ and $v_s = \pm v_r$, then

$$\begin{aligned}(\pi_\alpha^{\oplus j} \circ \eta \circ \psi_{\mathcal{M}})&(\lambda_1 v_1 + \cdots + \lambda_k v_k) = \\ &= (\pi_\alpha^{\oplus j} \circ \eta \circ \psi_{\mathcal{M}'})(\lambda_1 v_1 + \cdots + \lambda_k v_k) = (\pm\lambda, \dots, \pm\lambda),\end{aligned}$$

where $\eta\colon W_k \oplus U_k^{\oplus j} \longrightarrow U_k^{\oplus j}$ is the projection. Finally, in the case when $\lambda_s = 0$ for some $1 \leq s \leq k$, $\psi_{\mathcal{M}}(\lambda_1 v_1 + \cdots + \lambda_k v_k)$ and $\psi_{\mathcal{M}'}(\lambda_1 v_1 + \cdots + \lambda_k v_k)$ after projection to the sth coordinate of the subrepresentation W_k are equal to $-\frac{1}{k}$. □

Denote the radial projections by

$$\rho\colon U_k^{\oplus j}\backslash\{0\} \longrightarrow S(U_k^{\oplus j}) \quad \text{and} \quad \nu\colon (W_k \oplus U_k^{\oplus j})\backslash\{0\} \longrightarrow S(W_k \oplus U_k^{\oplus j}).$$

Note that ρ and ν are $\mathfrak{S}_k^{\pm}$-equivariant maps. Now the criterion stated in Proposition 2.1 (ii) can be strengthened as follows.

Theorem 2.3. *Let $d \geq 1$, $k \geq 1$ and $j \geq 1$ be integers, and let $\mathcal{M}$ be a collection of j masses in $\mathbb{R}^d$. We have the following two criteria:*

(i) *If there is no $\mathfrak{S}_k^{\pm}$-equivariant map*

$$Y_{d,k} \longrightarrow S(U_k^{\oplus j})$$

whose restriction to $Y_{d,k}^{>1}$ is $\mathfrak{S}_k^{\pm}$-homotopic to $\rho \circ \phi_{\mathcal{M}}|_{Y_{d,k}^{>1}}$, then $\Delta(j,k) \leq d$.

(ii) *If there is no $\mathfrak{S}_k^{\pm}$-equivariant map*

$$X_{d,k} \longrightarrow S(W_k \oplus U_k^{\oplus j})$$

whose restriction to $X_{d,k}^{>1}$ is $\mathfrak{S}_k^{\pm}$-homotopic to $\nu \circ \psi_{\mathcal{M}}|_{X_{d,k}^{>1}}$, then $\Delta(j,k) \leq d$.

2.6 Applying relative equivariant obstruction theory

In order to prove Theorems 1.4, 1.5, and 1.6 via Theorem 2.3(ii), we study the existence of an $\mathfrak{S}_k^{\pm}$-equivariant map

$$X_{d,k} \longrightarrow S(W_k \oplus U_k^{\oplus j}), \tag{4}$$

whose restriction to $X_{d,k}^{>1}$ is $\mathfrak{S}_k^{\pm}$-homotopic to $\nu \circ \psi_{\mathcal{M}}|_{X_{d,k}^{>1}}$ for some fixed collection $\mathcal{M}$ of j masses in $\mathbb{R}^d$. If we prove that such a map cannot exist, Theorems 1.4, 1.5, and 1.6 follow.

Denote by

$$N_1 := (d+1)k - 1$$

the dimension of the sphere $X_{d,k} = (S^d)^{*k}$, and by

$$N_2 := (2^k - 1)j + k - 2$$

the dimension of the sphere $S(W_k \oplus U_k^{\oplus j})$.
If $N_1 \le N_2$, then

$$\dim X_{d,k} = N_1 \le \operatorname{conn}\big(S(W_k \oplus U_k^{\oplus j})\big) + 1 = N_2.$$

Consequently, all obstructions to the existence of an $\mathfrak{S}_k^{\pm}$-equivariant map (4) vanish and so the map exists. Here $\operatorname{conn}(\cdot)$ denotes the connectivity of a space. Therefore, we assume that $N_1 > N_2$, which is equivalent to the Ramos lower bound $d \ge \frac{2^k-1}{k} j$. Furthermore, the following prerequisites for applying equivariant obstruction theory are satisfied:

- The N_1-sphere $X_{d,k}$ can be given the structure of a relative $\mathfrak{S}_k^{\pm}$-CW complex $X := (X_{d,k}, X_{d,k}^{>1})$ with a free $\mathfrak{S}_k^{\pm}$-action on $X_{d,k} \backslash X_{d,k}^{>1}$: In Section 3 we construct an explicit relative $\mathfrak{S}_k^{\pm}$-CW complex that models $X_{d,k}$.
- The sphere $S(W_k \oplus U_k^{\oplus j})$ is path connected and N_2-simple, except in the trivial case of $k = j = 1$ when $N_2 = 0$. Indeed, the group $\pi_1(S(W_k \oplus U_k^{\oplus j}))$ is abelian for $N_2 = 1$ and trivial for $N_2 > 1$ and therefore its action on $\pi_{N_2}(S(W_k \oplus U_k^{\oplus j}))$ is trivial.
- The $\mathfrak{S}_k^{\pm}$-equivariant map $h\colon X_{d,k}^{>1} \longrightarrow S(W_k \oplus U_k^{\oplus j})$ given by the composition $h := \nu \circ \psi_{\mathcal{M}}|_{X_{d,k}^{>1}}$, for a fixed collection of j masses $\mathcal{M}$, serves as the base map for extension.

Since the sphere $S(W_k \oplus U_k^{\oplus j})$ is (N_2-1)-connected, the map h can be extended to a $\mathfrak{S}_k^{\pm}$-equivariant map from the N_2-skeleton $X^{(N_2)} \longrightarrow S(W_k \oplus U_k^{\oplus j})$. A necessary criterion for the existence of the $\mathfrak{S}_k^{\pm}$-equivariant map (4) extending h is that the $\mathfrak{S}_k^{\pm}$-equivariant map $h = \nu \circ \psi_{\mathcal{M}}|_{X_{d,k}^{>1}}$ can be extended to a map from the (N_2+1)-skeleton $X^{(N_2+1)} \longrightarrow S(W_k \oplus U_k^{\oplus j})$.
Given the above hypotheses, we can apply relative equivariant obstruction theory, as presented by tom Dieck [7, Sec. II.3], to decide the existence of such an extension.
If g is an equivariant extension of h to the N_2-skeleton $X^{(N_2)}$, then the obstruction to extending g to the (N_2+1)-skeleton is encoded by the equivariant cocycle

$$\mathfrak{o}(g) \in \mathcal{C}_{\mathfrak{S}_k^{\pm}}^{N_2+1}\big(X_{d,k}, X_{d,k}^{>1}; \pi_{N_2}(S(W_k \oplus U_k^{\oplus j}))\big).$$

The $\mathfrak{S}_k^{\pm}$-equivariant map $g\colon X^{(N_2)} \longrightarrow S(W_k \oplus U_k^{\oplus j})$ extends to $X^{(N_2+1)}$ if and only if $\mathfrak{o}(g) = 0$. Furthermore, the cohomology class

$$[\mathfrak{o}(g)] \in \mathcal{H}_{\mathfrak{S}_k^{\pm}}^{N_2+1}\big(X_{d,k}, X_{d,k}^{>1}; \pi_{N_2}(S(W_k \oplus U_k^{\oplus j}))\big),$$

vanishes if and only if the restriction $g|_{X^{(N_2-1)}}$ to the (N_2-1)-skeleton can be extended to the (N_2+1)-skeleton $X^{(N_2+1)}$. Any two extensions g and g' of h to the N_2-skeleton are equivariantly homotopic on the (N_2-1)-skeleton and therefore the cohomology classes coincide: $[\mathfrak{o}(g)] = [\mathfrak{o}(g')]$. Hence, it suffices to compute the cohomology class $[\mathfrak{o}(\nu \circ \psi_{\mathcal{M}}|_{X^{(N_2)}})]$ for a fixed collection of j masses $\mathcal{M}$ with the property that $0 \notin \operatorname{im}(\psi_{\mathcal{M}}|_{X^{(N_2)}})$.
Let f be the attaching map for an (N_2+1)-cell θ and e its corresponding basis element in the cellular chain group $C_{N_2+1}(X_{d,k}, X_{d,k}^{>1})$. Then

$$\mathfrak{o}(\nu \circ \psi_{\mathcal{M}}|_{X^{(N_2)}})(e) = [\nu \circ \psi_{\mathcal{M}} \circ f|_{\partial\theta}]$$

is the homotopy class of the map represented by the composition

$$\partial\theta_j \xrightarrow{f|_{\partial\theta}} X^{(N_2)} \xrightarrow{\nu\circ\psi_{\mathcal{M}}|_{X^{(N_2)}}} S(W_k \oplus U_k^{\oplus j}).$$

Since $\partial\theta$ and $S(W_k \oplus U_k^{\oplus j})$ are spheres of the same dimension N_2, the homotopy class $[\nu \circ \psi_{\mathcal{M}} \circ f|_{\partial\theta}]$ is determined by the degree of the map $\nu \circ \psi_{\mathcal{M}} \circ f|_{\partial\theta}$. Here we assume that the $\mathfrak{S}_k^{\pm}$-CW structure on $X_{d,k}$ is endowed with cell orientations, and in addition an orientation on the sphere $S(W_k \oplus U_k^{\oplus j})$ is fixed in advance. Therefore, the degree of the map $\nu \circ \psi_{\mathcal{M}} \circ f|_{\partial\theta}$ is well-defined.
Let $\alpha := \psi_{\mathcal{M}} \circ f|_{\partial\theta}$. In order to compute the degree of the map $\nu \circ \alpha$ and consequently the obstruction cocycle evaluated at e, fix the collection of measures as follows. Let $\mathcal{M}$ be the collection of masses $(I_1, \dots, I_j)$ where I_r is the mass concentrated on the segment $\gamma((t_r^1, t_r^2))$ of the moment curve in $\mathbb{R}^d$

$$\gamma(t) = (t, \tbinom{t}{2}, \tbinom{t}{3}, \dots, \tbinom{t}{d})^t,$$

such that

$$\ell < t_1^1 < t_1^2 < t_2^1 < t_2^2 < \dots < t_j^1 < t_j^2,$$

for an integer ℓ, $0 \le \ell \le d-1$. The intervals $(I_1, \dots, I_j)$ determined by numbers $t_r^1 < t_r^2$ can be chosen in such a way that $0 \notin \operatorname{im}(\psi_{\mathcal{M}}|_{X^{(N_2)}})$. For every concrete situation in Section 4 this is verified directly.
Now consider the following commutative diagram:

$$\begin{tikzcd}
\partial\theta \arrow[r, "f|_{\partial\theta}"] \arrow[d] & X^{(N_2)} \arrow[r, "\psi_{\mathcal{M}}|_{X^{(N_2)}}"] \arrow[d] & W_k \oplus U_k^{\oplus j}\backslash\{0\} \arrow[r, "\nu"] \arrow[d] & S(W_k \oplus U_k^{\oplus j}) \\
\theta \arrow[r, "f"] & X^{(N_2+1)} \arrow[r, "\psi_{\mathcal{M}}|_{X^{(N_2+1)}}"] & W_k \oplus U_k^{\oplus j} &
\end{tikzcd}$$

where the vertical arrows are inclusions, and the composition of the lower horizontal maps is denoted by $\beta := \psi_{\mathcal{M}}|_{X^{(N_2+1)}} \circ f$. Furthermore, let $B_\varepsilon(0)$ be a ball with center 0 in $W_k \oplus U_k^{\oplus j}$ of sufficiently small radius $\varepsilon > 0$. Set $\widetilde{\theta} := \theta \backslash \beta^{-1}(B_\varepsilon(0))$. Since $\dim\theta = \dim W_k \oplus U_k^{\oplus j}$ we can assume that the set of zeros $\beta^{-1}(0) \subset \operatorname{relint}\theta$ is finite, say of cardinality $r \ge 0$. Again, in every calculation presented in Section 4 this assumption is explicitly verified. The function β is a restriction of the test map and therefore *the points in* $\beta^{-1}(0)$ *correspond to arrangements of* k *hyperplanes* $\mathcal{H}$ *in* $\operatorname{relint}\theta$ *that equipart* $\mathcal{M}$. Moreover, the facts that the measures are intervals on a moment curve and that each hyperplane of the arrangement from $\beta^{-1}(0)$ cuts the moment curve in d distinct points imply that each zero in $\beta^{-1}(0)$ is isolated and transversal. The boundary of $\widetilde{\theta}$ consists of the boundary $\partial\theta$ and r disjoint copies of N_2-spheres $S_1,\dots,S_r$, one for each zero of β on θ. Consequently, the fundamental class of $\partial\theta$ is equal to the sum of fundamental classes $\sum[S_i]$ in $H_{N_1}(\widetilde{\theta};\mathbb{Z})$. Here the fundamental class of $\partial\theta$ is determined by the cell orientation inherited from the $\mathfrak{S}_k^{\pm}$-CW structure on $X_{d,k}$. The fundamental classes of $[S_i]$ are determined in such a way that the equality $[\partial\theta]=\sum[S_i]$ holds. Thus

$$\sum(\nu\circ\beta|_{\widetilde{\theta}})_*([S_i])=(\nu\circ\beta|_{\widetilde{\theta}})_*([\partial\theta])=(\nu\circ\alpha)_*([\partial\theta])=\deg(\nu\circ\alpha)\cdot[S(W_k\oplus U_k^{\oplus j})].$$

Recall, we have fixed the orientation on the sphere $S(W_k\oplus U_k^{\oplus j})$ and so the fundamental class $[S(W_k\oplus U_k^{\oplus j})]$ is also completely determined. On the other hand,

$$\sum(\nu\circ\beta|_{S_i})_*([S_i])=\Big(\sum\deg(\nu\circ\beta|_{S_i})\Big)\cdot[S(W_k\oplus U_k^{\oplus j})].$$

Hence, $\deg(\nu\circ\alpha)=\sum\deg(\nu\circ\beta|_{S_i})$ where the sum ranges over all arrangements of k hyperplanes $\mathcal{H}$ in $\operatorname{relint}\theta$ that equipart $\mathcal{M}$; consult [14, Prop. IV.4.5]. In other words,

$$\mathfrak{o}(\nu\circ\psi_{\mathcal{M}}|_{X^{(N_2)}})(e)=[\nu\circ\psi_{\mathcal{M}}\circ f|_{\partial\theta}]=\deg(\nu\circ\alpha)\cdot\zeta=\sum\deg(\nu\circ\beta|_{S_i})\cdot\zeta,\quad(5)$$

where $\zeta\in\pi_{N_2}(S(W_k\oplus U_k^{\oplus j}))\cong\mathbb{Z}$ is a generator, and the sum ranges over all arrangements of k hyperplanes $\mathcal{H}$ in $\operatorname{relint}\theta$ that equipart $\mathcal{M}$.

If in addition we assume that all local degrees $\deg(\nu\circ\beta|_{S_i})$ are ±1 and that the number of arrangements of k hyperplanes $\mathcal{H}$ in $\operatorname{relint}\theta$ that equipart $\mathcal{M}$ is odd, then we conclude that $\mathfrak{o}(\nu\circ\psi_{\mathcal{M}}|_{X^{(N_2)}})(e)\neq0$. It will turn out that in many situations this information implies that the cohomology class $[\mathfrak{o}(\nu\circ\psi_{\mathcal{M}})]$ is not zero, and consequently the related $\mathfrak{S}_k^{\pm}$-equivariant map (4) does not exist, concluding the proof of corresponding Theorems 1.4, 1.5, and 1.6.

3 A regular cell complex model for the join configuration space

In this section, motivated by methods used in [2] and [6], we construct a regular $\mathfrak{S}_k^{\pm}$-CW model for the join configuration space $X_{d,k}=(S^d)^{*k}\cong S(\mathbb{R}^{(d+1)\times k})$ such that $X_{d,k}^{>1}$ is a $\mathfrak{S}_k^{\pm}$-CW subcomplex. Consequently, $(X_{d,k},X_{d,k}^{>1})$ has the structure of a relative $\mathfrak{S}_k^{\pm}$-CW complex. For simplicity the cell complex we construct is denoted by $X:=(X_{d,k},X_{d,k}^{>1})$ as well. The cell model is obtained in two steps:

(1) the vector space $\mathbb{R}^{(d+1)\times k}$ is decomposed into a union of disjoint relatively open cones (each containing the origin in its closure) on which the $\mathfrak{S}_k^{\pm}$-action operates linearly permuting the cones, and then
(2) the open cells of a regular $\mathfrak{S}_k^{\pm}$-CW model are obtained as intersections of these relatively open cones with the unit sphere $S(\mathbb{R}^{(d+1)\times k})$.

The explicit relative $\mathfrak{S}_k^{\pm}$-CW complex we construct here is an essential object needed for the study of the existence of $\mathfrak{S}_k^{\pm}$-equivariant maps $X_{d,k} \longrightarrow S(W_k \oplus U_k^{\oplus j})$ via the relative equivariant obstruction theory of tom Dieck [7, Sec. II.3].

3.1 Stratifications by cones associated to an arrangement

The first step in the construction of the $\mathfrak{S}_k^{\pm}$-CW model is an appropriate stratification of the ambient space $\mathbb{R}^{(d+1)\times k}$. First we introduce the notion of the stratification of a Euclidean space and collect some relevant properties.

Definition 3.1. Let $\mathcal{A}$ be an arrangement of linear subspaces in a Euclidean space E. A *stratification of E (by cones) associated to $\mathcal{A}$* is a finite collection $\mathcal{C}$ of subsets of E that satisfies the following properties:

(i) $\mathcal{C}$ consists of finitely many non-empty relatively open polyhedral cones in E.
(ii) $\mathcal{C}$ is a partition of E, i.e., $E = \biguplus_{C\in\mathcal{C}} C$.
(iii) The closure $\overline{C}$ of every cone $C \in \mathcal{C}$ is a union of cones in $\mathcal{C}$.
(iv) Every subspace $A \in \mathcal{A}$ is a union of cones in $\mathcal{C}$.

An element of the family $\mathcal{C}$ is called a *stratum*.

Example 3.2. Let E be a Euclidean space of dimension d, let L be a linear subspace of codimension r, where $1 \le r \le d$, and let $\mathcal{A}$ be the arrangement $\{L\}$. Choose a flag that terminates at L, i.e., fix a sequence of linear subspaces in E

$$E = L^{(0)} \supset L^{(1)} \supset \cdots \supset L^{(r)} = L, \tag{6}$$

so that $\dim L^{(i)} = d - i$. The family $\mathcal{C}$ associated to the flag (6) consists of L and of the connected components of the successive complements

$$L^{(0)}\backslash L^{(1)},\ L^{(1)}\backslash L^{(2)},\ \dots\ , L^{(r-1)}\backslash L^{(r)}.$$

A $L^{(i)}$ is a hyperplane in $L^{(i-1)}$, each of the complements $L^{(i-1)}\backslash L^{(i)}$ has two connected components. This indeed yields a stratification by cones for the arrangement $\mathcal{A}$ in E.

Definition 3.3. Let $(\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n)$ be a collection of arrangements of linear subspaces in the Euclidean space E and let $(\mathcal{C}_1, \mathcal{C}_2 \dots, \mathcal{C}_n)$ be the associated collection of stratifications of E by cones. The *common refinement* of the stratifications is the family

$$\mathcal{C} := \{C_1 \cap C_2 \cap \cdots \cap C_n \neq \emptyset : C_i \in \mathcal{C}_i \text{ for all } i\}.$$

In order to verify that the common refinement of stratifications is again a stratification, we use the following elementary lemma.

Lemma 3.4. *Let $A_1, \dots, A_n$ be relatively open convex sets in E that have non-empty intersection, $A_1 \cap \dots \cap A_n \neq \emptyset$. Then the following relation holds for the closures:*

$$\overline{A_1 \cap \dots \cap A_n} = \overline{A_1} \cap \dots \cap \overline{A_n}.$$

Proof. The inclusion "$\subseteq$" follows directly. For the opposite inclusion take $x \in \overline{A_1} \cap \dots \cap \overline{A_n}$. Choose a point $y \in A_1 \cap \dots \cap A_n \neq \emptyset$ and consider the line segment $(x, y] := \{\lambda x + (1-\lambda)y : 0 \leq \lambda < 1\}$. As each A_i is relatively open, the segment $(x, y]$ is contained in each of the A_i and consequently it is contained in $A_1 \cap \dots \cap A_n$. Thus we obtain a sequence in this intersection converging to x, which implies that $x \in \overline{A_1 \cap \dots \cap A_n}$. □

Proposition 3.5. *Given stratifications by cones $\mathcal{C}_1, \mathcal{C}_2 \dots, \mathcal{C}_n$ associated to linear subspace arrangements $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n$, their common refinement is a stratification by cones associated to the subspace arrangement $\mathcal{A} := \mathcal{A}_1 \cup \dots \cup \mathcal{A}_n$.*

Proof. Properties (i) and (ii) of Definition 3.1 follow immediately from the definition of the common refinement. To verify property (iv), observe that a subspace $A_t \in \mathcal{A}_t$ is a union of strata from $\mathcal{C}_t$, say $A_t = \bigcup_s U_{t,s}$ where $U_{t,s} \in \mathcal{C}_t$. Hence, taking the union of intersections $C_1 \cap \dots \cap U_{t,s} \cap \dots \cap C_n$ for all $C_i \in \mathcal{C}_i$ where $i \neq t$, and all $U_{t,s}$ gives A_t. Property (iii) follows from Lemma 3.4. □

Example 3.6. Let E be a Euclidean space of dimension d and let $\mathcal{A} = \{L_1, \dots, L_s\}$ be an arrangement of linear subspaces of E. As in Example 3.2, for each of the subspaces L_i in the arrangement $\mathcal{A}$ fix a flag $L_i^{(s)}$ and form the corresponding stratifications $\mathcal{C}_1, \dots, \mathcal{C}_s$. The common refinement of stratifications $\mathcal{C}_1, \dots, \mathcal{C}_s$ is a stratification by cones associated to the subspace arrangement $\mathcal{A}$.

An arrangement of linear subspaces is *essential* if the intersection of the subspaces in the arrangement is $\{0\}$.

Proposition 3.7. *The intersection of a stratification $\mathcal{C}$ of E by cones associated to an essential linear subspace arrangement with the sphere $S(E)$ gives a regular CW-complex.*

Proof. The elements $C \in \mathcal{C}$ are relative open polyhedral cones. As $\{0\}$ is a stratum by itself, none of the strata contains a line through the origin. Thus $C \cap S(E)$ is an open cell, whose closure $\overline{C} \cap S(E)$ is a finite union of cells of the form $C' \cap S(E)$, so we get a regular CW complex. □

3.2 A stratification of $\mathbb{R}^{(d+1)\times k}$

Now we introduce the stratification of $\mathbb{R}^{(d+1)\times k}$ that will give us a $\mathfrak{S}_k^{\pm}$-CW model for $X_{d,k}$. One version of it, $\mathcal{C}$, arises from the construction in the previous section. However, we also give combinatorial descriptions of relatively-open convex cones in the stratification $\mathcal{C}'$ directly, for which the action of $\mathfrak{S}_k^{\pm}$ is evident. We then verify that $\mathcal{C}$ and $\mathcal{C}'$ coincide.

3.2.1 Stratification

Let elements $x \in \mathbb{R}^{(d+1)\times k}$ be written as $x = (x_1, \dots, x_k)$ where $x_i = (x_{t,i})_{t\in[d+1]}$ is the i-th column of the matrix x. Consider the arrangement $\mathcal{A}$ consisting of the following subspaces:

$$\begin{aligned}
L_r &:= \{(x_1,\dots,x_k) \in \mathbb{R}^{(d+1)\times k} : x_r = 0\}, && 1 \le r \le k \\
L_{r,s}^+ &:= \{(x_1,\dots,x_k) \in \mathbb{R}^{(d+1)\times k} : x_r - x_s = 0\}, && 1 \le r < s \le k \\
L_{r,s}^- &:= \{(x_1,\dots,x_k) \in \mathbb{R}^{(d+1)\times k} : x_r + x_s = 0\}, && 1 \le r < s \le k.
\end{aligned}$$

With each subspace we associate a flag:

(i) With $L_r = \{x_r = 0\}$ we associate

$$\mathbb{R}^{(d+1)\times k} \supset \{x_{1,r} = 0\} \supset \{x_{1,r} = x_{2,r} = 0\} \supset \cdots \supset \{x_{1,r} = x_{2,r} = \cdots = x_{d+1,r} = 0\},$$

(ii) With $L_{r,s}^+ = \{x_r - x_s = 0\}$ we associate

$$\mathbb{R}^{(d+1)\times k} \supset \{x_{1,r} - x_{1,s} = 0\} \supset \{x_{1,r} - x_{1,s} = x_{2,r} - x_{2,s} = 0\} \supset \cdots \supset \{x_{1,r} - x_{1,s} = x_{2,r} - x_{2,s} = \cdots = x_{d+1,r} - x_{d+1,s} = 0\},$$

(iii) $L_{r,s}^- = \{x_r + x_s = 0\}$ we associate

$$\mathbb{R}^{(d+1)\times k} \supset \{x_{1,r} + x_{1,s} = 0\} \supset \{x_{1,r} + x_{1,s} = x_{2,r} + x_{2,s} = 0\} \supset \cdots \supset \{x_{1,r} + x_{1,s} = x_{2,r} + x_{2,s} = \cdots = x_{d+1,r} + x_{d+1,s} = 0\}.$$

The construction from Example 3.2 shows how every subspace in $\mathcal{A}$ leads to a stratification by cones of $\mathbb{R}^{(d+1)\times k}$. The stratifications associated to the subspaces $L_r, L_{r,s}^+, L_{r,s}^-$ are denoted by $\mathcal{C}_r, \mathcal{C}_{r,s}^+, \mathcal{C}_{r,s}^-$, respectively. Now, if we apply Example 3.6 to this concrete situation we obtain the stratification by cones $\mathcal{C}$ of $\mathbb{R}^{(d+1)\times k}$ associated to the subspace arrangement $\mathcal{A}$. This means that each stratum of $\mathcal{C}$ is a non-empty intersection of strata from the stratifications $\mathcal{C}_r, \mathcal{C}_{r,s}^+, \mathcal{C}_{r,s}^-$ where $1 \le r < s \le k$.

3.2.2 Partition

Let us fix:

- a permutation $\sigma := (\sigma_1, \sigma_2, \dots, \sigma_k) \equiv (\sigma_1\sigma_2\dots\sigma_k) \in \mathfrak{S}_k$, $\sigma\colon t \mapsto \sigma_t$,
- a collection of signs $S := (s_1, \dots, s_k) \in \{+1, -1\}^k$, and
- integers $I := (i_1, \dots, i_k) \in \{1, \dots, d+2\}^k$.

Furthermore, define x_0 to be the origin in $\mathbb{R}^{(d+1)\times k}$, $\sigma_0 = 0$ and $s_0 = 1$. Define

$$C_I^S(\sigma) = C_{i_1,\dots,i_k}^{s_1,\dots,s_k}(\sigma_1,\sigma_2,\dots,\sigma_k) \subseteq \mathbb{R}^{(d+1)\times k}$$

to be the set of all points $(x_1,\dots,x_k) \in \mathbb{R}^{(d+1)\times k}$, $x_i = (x_{1,i},\dots,x_{d+1,i})$, such that for each $1 \le t \le k$,

- if $1 \le i_t \le d+1$, then $s_{t-1}x_{i_t,\sigma_{t-1}} < s_t x_{i_t,\sigma_t}$ with $s_{t-1}x_{i',\sigma_{t-1}} = s_t x_{i',\sigma_t}$ for every $i' < i_t$,
- if $i_t = d+2$, then $s_{i_{t-1}}x_{\sigma_{t-1}} = s_{i_t}x_{\sigma_t}$.

Any triple $(\sigma|I|S) \in \mathfrak{S}_k \times \{1,\dots,d+2\}^k \times \{+1,-1\}^k$ is called a *symbol.* In the notation of symbols we abbreviate signs $\{+1,-1\}$ by $\{+,-\}$. The defining set of "inequalities" for the stratum $C_I^S(\sigma)$ is briefly denoted by:

$$\begin{aligned} C_I^S(\sigma) &= C_{i_1,\dots,i_k}^{s_1,\dots,s_k}(\sigma_1,\sigma_2,\dots,\sigma_k) \\ &= \{(x_1,\dots,x_k) \in \mathbb{R}^{(d+1)\times k} : 0 <_{i_1} s_1x_{\sigma_1} <_{i_2} s_2x_{\sigma_2} <_{i_3} \cdots <_{i_k} s_kx_{\sigma_k}\}, \end{aligned}$$

where $y <_i y'$, for $1 \le i \le d+1$, means that y and y' agree in the first $i-1$ coordinates and at the i-th coordinate $y_i < y'_i$. The inequality $y <_{d+2} y'$ denotes that $y = y'$. Each set $C_I^S(\sigma)$ is the relative interior of a polyhedral cone in $(\mathbb{R}^{d+1})^k$ of codimension $(i_1-1)+\cdots+(i_k-1)$, i.e.,

$$\dim C_{i_1,\dots,i_k}^{s_1,\dots,s_k}(\sigma_1,\sigma_2,\dots,\sigma_k) = (d+2)k - (i_1+\cdots+i_k).$$

Let $\mathcal{C}'$ denote the family of strata $C_I^S(\sigma)$ defined by all symbols, i.e.,

$$\mathcal{C}' = \{C_I^S(\sigma) : (\sigma|I|S) \in \mathfrak{S}_k \times \{1,\dots,d+2\}^k \times \{+1,-1\}^k\}.$$

Different symbols can define the same set, and

$$C_I^S(\sigma) \cap C_{I'}^{S'}(\sigma) \ne \emptyset \iff C_I^S(\sigma) = C_{I'}^{S'}(\sigma).$$

In order to verify that the family $\mathcal{C}'$ is a partition of $\mathbb{R}^{(d+1)\times k}$ it remains to prove that it is a covering.

Lemma 3.8. $\bigcup \mathcal{C}' = \mathbb{R}^{(d+1)\times k}$.

Proof. Let $(x_1,\dots,x_k) \in \mathbb{R}^{(d+1)\times k}$. First, choose signs $r_1,\dots,r_k \in \{+1,-1\}$ so that the vectors $r_1x_1,\dots,r_kx_k$ are greater or equal to $0 \in \mathbb{R}^{(d+1)\times k}$ with respect to the lexicographic order, i.e., the first non-zero coordinate of each of the vectors r_ix_i is greater than zero. The choice of signs is not unique if one of the vectors x_i is zero. Next, record a permutation $\sigma \in \mathfrak{S}_k$ such that

$$0 <_{\text{lex}} r_{\sigma_1}x_{\sigma_1} <_{\text{lex}} r_{\sigma_2}x_{\sigma_2} <_{\text{lex}} \cdots <_{\text{lex}} r_{\sigma_k}x_{\sigma_k},$$

where $<_{\text{lex}}$ denotes the lexicographic order. The permutation σ is not unique if $r_ix_i = r_tx_t$ for some $i \ne t$. Define $s_i := r_{\sigma_i}$. Finally, collect coordinates i_t where vectors $s_{t-1}x_{\sigma_{t-1}}$ and $s_tx_{\sigma_t}$ first differ, or put $i_t = d+2$ if they coincide. Thus, $(x_1,\dots,x_k) \in C_{i_1,\dots,i_k}^{s_1,\dots,s_k}(\sigma_1,\sigma_2,\dots,\sigma_k)$. $\square$

Example 3.9. Let $d = 0$ and $k = 2$. Then the plane $\mathbb{R}^2$ is decomposed into the following cones. There are 8 open cones of dimension 2:

$$\begin{array}{rcl}
C^{+;+}_{1,1}(12) & = & \{(x_1, x_2) \in \mathbb{R}^2 : 0 < x_1 < x_2\}, \\
C^{-;+}_{1,1}(12) & = & \{(x_1, x_2) \in \mathbb{R}^2 : 0 < -x_1 < x_2\}, \\
C^{+;-}_{1,1}(12) & = & \{(x_1, x_2) \in \mathbb{R}^2 : 0 < x_1 < -x_2\}, \\
C^{-;-}_{1,1}(12) & = & \{(x_1, x_2) \in \mathbb{R}^2 : 0 < -x_1 < -x_2\}, \\
C^{+;+}_{1,1}(21) & = & \{(x_1, x_2) \in \mathbb{R}^2 : 0 < x_2 < x_1\}, \\
C^{-;+}_{1,1}(21) & = & \{(x_1, x_2) \in \mathbb{R}^2 : 0 < -x_2 < x_1\}, \\
C^{+;-}_{1,1}(21) & = & \{(x_1, x_2) \in \mathbb{R}^2 : 0 < x_2 < -x_1\}, \\
C^{-;-}_{1,1}(21) & = & \{(x_1, x_2) \in \mathbb{R}^2 : 0 < -x_2 < -x_1\}.
\end{array}$$

Furthermore, there are 8 cones of dimension 1:

$$\begin{array}{rcccl}
C^{+;+}_{1,2}(12) & = & C^{+;+}_{1,2}(21) & = & \{(x_1, x_2) \in \mathbb{R}^2 : 0 < x_1 = x_2\}, \\
C^{-;+}_{1,2}(12) & = & C^{+;-}_{1,2}(21) & = & \{(x_1, x_2) \in \mathbb{R}^2 : 0 < -x_1 = x_2\}, \\
C^{+;-}_{1,2}(12) & = & C^{-;+}_{1,2}(21) & = & \{(x_1, x_2) \in \mathbb{R}^2 : 0 < x_1 = -x_2\}, \\
C^{-;-}_{1,2}(12) & = & C^{-;-}_{1,2}(21) & = & \{(x_1, x_2) \in \mathbb{R}^2 : 0 < -x_1 = -x_2\}, \\
C^{+;+}_{2,1}(12) & = & C^{-;+}_{2,1}(12) & = & \{(x_1, x_2) \in \mathbb{R}^2 : 0 = x_1 < x_2\}, \\
C^{+;-}_{2,1}(12) & = & C^{-;-}_{2,1}(12) & = & \{(x_1, x_2) \in \mathbb{R}^2 : 0 = x_1 < -x_2\}, \\
C^{+;+}_{2,1}(21) & = & C^{-;+}_{2,1}(21) & = & \{(x_1, x_2) \in \mathbb{R}^2 : 0 = x_2 < x_1\}, \\
C^{+;-}_{2,1}(21) & = & C^{-;-}_{2,1}(21) & = & \{(x_1, x_2) \in \mathbb{R}^2 : 0 = x_2 < -x_1\}.
\end{array}$$

The origin in $\mathbb{R}^2$ is given by $C^{\pm;\pm}_{2,2}(12) = C^{\pm;\pm}_{2,2}(21)$. The example illustrates a property of our decomposition of $\mathbb{R}^{(d+1)\times k}$: There is a surjection from symbals to cones that is not a bijection, i.e., different symbols can define the same cones.

Example 3.10. Let $d = 2$ and $k = 4$. The stratum associated to the symbol $(2143 \,|\, 2, 3, 1, 4 \,|\, +1, -1, +1, -1)$ can be described explicitly as follows.

$$\left\{ \begin{pmatrix} x_{1,1} & x_{1,2} & x_{1,3} & x_{1,4} \\ x_{2,1} & x_{2,2} & x_{2,3} & x_{2,4} \\ x_{3,1} & x_{3,2} & x_{3,3} & x_{3,4} \end{pmatrix} \in (\mathbb{R}^3)^4 : \right.$$

$$\left. \begin{array}{ccccccccc}
0 & = & x_{1,2} & = & -x_{1,1} & < & x_{1,4} & = & -x_{1,3} \\
0 & < & x_{2,2} & = & -x_{2,1} & & x_{2,4} & = & -x_{2,3} \\
 & & x_{3,2} & < & -x_{3,1} & & x_{3,4} & = & -x_{3,3}
\end{array} \right\}.$$

In particular,

$$C^{+,-,+,-}_{2,3,1,4}(2143) = C^{+,-,-,+}_{2,3,1,4}(2134).$$

3.2.3 $\mathcal{C}$ and $\mathcal{C}'$ coincide

We proved that $\mathcal{C}$ is a stratification by cones of $\mathbb{R}^{(d+1)\times k}$, and that $\mathcal{C}'$ is a partition of $\mathbb{R}^{(d+1)\times k}$. Since both $\mathcal{C}$ and $\mathcal{C}'$ are partitions it suffices to prove

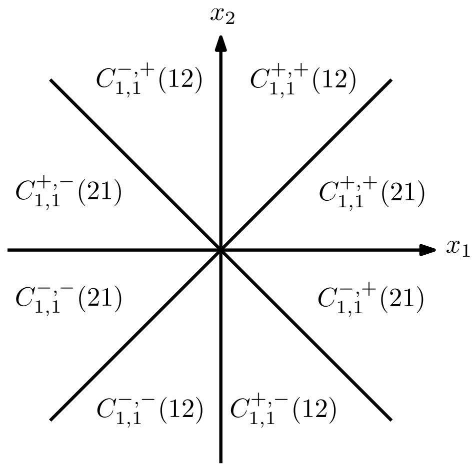


Figure 1: Illustration of the stratification in Example 3.9

that for every symbol $(\sigma|I|S) \in \mathfrak{S}_k \times \{1, \dots, d+2\}^k \times \{+1, -1\}^k$ the cone $C_I^S(\sigma) \in \mathcal{C}'$ also belongs to $\mathcal{C}$.
Consider the cone $C_I^S(\sigma)$ in $\mathcal{C}'$. It is determined by

$$\begin{aligned} C_I^S(\sigma) &= C_{i_1,\dots,i_k}^{s_1,\dots,s_k}(\sigma_1, \sigma_2, \dots, \sigma_k) \\ &= \{(x_1, \dots, x_k) \in \mathbb{R}^{(d+1)\times k} : 0 <_{i_1} s_1 x_{\sigma_1} <_{i_2} s_2 x_{\sigma_2} <_{i_3} \cdots <_{i_k} s_k x_{\sigma_k}\}. \end{aligned}$$

The defining inequalities for $C_I^S(\sigma)$ imply that $(x_1, \dots, x_k) \in C_I^S(\sigma)$ if and only if

- $0 <_{\min\{i_1,\dots,i_a\}} s_a x_a$ for $1 \le a \le k$, and
- $s_a x_a <_{\min\{i_{a+1},\dots,i_b\}} s_b x_b$ for $1 \le a < b \le k$,

if and only if

- $(x_1, \dots, x_k)$ belongs to the appropriate one of two strata in the complement

$$L_a^{(\min\{i_1,\dots,i_a\}-1)} \setminus L_a^{(\min\{i_1,\dots,i_a\}-2)}$$

  of the stratification $\mathcal{C}_a$ depending on the sign s_a where $1 \le a \le k$, and
- $(x_1, \dots, x_k)$ belongs to the appropriate one of two strata in the complement

$$L_{a,b}^{s_a s_b (\min\{i_{a+1},\dots,i_b\}-1)} \setminus L_{a,b}^{s_a s_b (\min\{i_{a+1},\dots,i_b\}-2)}$$

  of the stratification $\mathcal{C}_{a,b}^{s_a s_b}$ depending on the sign of the product $s_a s_b$ where $1 \le a < b \le k$. The product $s_a s_b$, appearing in the “exponent notation” of $L_{a,b}^{s_a s_b}$, is either “+” when the product $s_a s_b = 1$, or “−” when $s_a s_b = -1$.

Here we use the notation of Examples 3.2 and 3.6.
Thus we have proved that $C_I^S(\sigma) \in \mathcal{C}$ and consequently $\mathcal{C} = \mathcal{C}'$.

3.3 The $\mathfrak{S}_k^{\pm}$-CW model for $X_{d,k}$

The action of the group $\mathfrak{S}_k^{\pm}$ on the space $\mathbb{R}^{(d+1)\times k}$ induces an action on the family of strata $\mathcal{C}$ by as follows:

$$\begin{aligned} \pi \cdot C_I^S(\sigma) &= C_I^S(\pi\sigma), & (7)\\ \varepsilon_t \cdot C_I^S(\sigma) &= \varepsilon_t \cdot C_{i_1,\dots,i_k}^{s_1,\dots,s_k}(\sigma_1,\sigma_2,\dots,\sigma_k) \\ &= C_{i_1,\dots,i_k}^{s_1,\dots,-s_t,\dots,s_k}(\sigma_1,\sigma_2,\dots,\sigma_k), & (8)\end{aligned}$$

where $\pi \in \mathfrak{S}_k$ and $\varepsilon_1,\dots,\varepsilon_k$ are the canonical generators of the subgroup $(\mathbb{Z}/2)^k$ of $\mathfrak{S}_k^{\pm}$.

The $\mathfrak{S}_k^{\pm}$-CW complex that models $X_{d,k} = S(\mathbb{R}^{(d+1)\times k})$ is obtained by intersecting each stratum $C_I^S(\sigma)$ with the unit sphere $S(\mathbb{R}^{(d+1)\times k})$. Each stratum is a relatively open cone that does not contain a line. Therefore the intersection

$$D_I^S(\sigma) = D_{i_1,\dots,i_k}^{s_1,\dots,s_k}(\sigma_1,\sigma_2,\dots,\sigma_k) := C_{i_1,\dots,i_k}^{s_1,\dots,s_k}(\sigma_1,\sigma_2,\dots,\sigma_k) \cap S(\mathbb{R}^{(d+1)\times k})$$

is an open cell of dimension $(d+2)k-(i_1+\dots+i_k)-1$. The action of $\mathfrak{S}_k^{\pm}$ is induced by (7) and (8):

$$\begin{aligned} \pi \cdot D_I^S(\sigma) &= D_I^S(\pi\sigma), & (9)\\ \varepsilon_t \cdot D_I^S(\sigma) &= \varepsilon_t \cdot D_{i_1,\dots,i_k}^{s_1,\dots,s_k}(\sigma_1,\sigma_2,\dots,\sigma_k) \\ &= D_{i_1,\dots,i_k}^{s_1,\dots,-s_t,\dots,s_k}(\sigma_1,\sigma_2,\dots,\sigma_k). & (10)\end{aligned}$$

Thus we have obtained a regular $\mathfrak{S}_k^{\pm}$-CW model for $X_{d,k}$. In particular, the action of the group $\mathfrak{S}_k^{\pm}$ on the space $\mathbb{R}^{(d+1)\times k}$ induces a cellular action on the model.

Theorem 3.11. *Let $d \geq 1$ and $k \geq 1$ be integers, and $N_1 = (d+1)k-1$. The family of cells*

$$\{D_I^S(\sigma) : (\sigma|I|S) \neq (\sigma|d+2,\dots,d+2|S)\}$$

forms a finite regular N_1-dimensional $\mathfrak{S}_k^{\pm}$-CW complex $X := (X_{d,k}, X_{d,k}^{>1})$ that models the join configuration space $X_{d,k} = S(\mathbb{R}^{(d+1)\times k})$. It has

- *one full $\mathfrak{S}_k^{\pm}$-orbit in maximal dimension N_1, and*
- *k full $\mathfrak{S}_k^{\pm}$-orbits in dimension $N_1 - 1$.*

The (cellular) $\mathfrak{S}_k^{\pm}$-action on $X_{d,k}$ is given by (9) and (10). Furthermore the collection of cells

$$\{D_I^S(\sigma) : i_s = d+2 \text{ for some } 1 \leq s \leq k\}$$

is a $\mathfrak{S}_k^{\pm}$-CW subcomplex and models $X_{d,k}^{>1}$.

Example 3.12. Let $d \geq 1$ and $k \geq 2$ be integers with $dk = (2^k-1)j + \ell$, where $0 \leq \ell \leq d-1$. Consider the cell $\theta := D_{\ell+1,1,1,\dots,1}^{+,+,+,\dots,+}(1,2,3,\dots,k)$ of dimension $N_1 - \ell = N_2 + 1$ in $X_{d,k}$. It is determined by the following inequalities:

$$0 <_{\ell+1} x_1 <_1 x_2 <_1 \cdots <_1 x_k.$$

For the process of determining the boundary of θ, depending on value of ℓ, we distinguish the following cases.

(1) Let $\ell = 0$. Then $\theta := D^{+,+,+,\dots,+}_{1,1,1,\dots,1}(1,2,3,\dots,k)$. The cells of codimension 1 in the boundary of θ are obtained by introducing one of the following extra equalities:

$$x_{1,1} = 0\,, \quad x_{1,1} = x_{1,2}\,, \quad \dots \quad x_{1,k-1} = x_{1,k}.$$

Each of these equalities will give two cells of dimension N_2, hence in total $2k$ cells of codimension 1, in the boundary of θ.

(A) The equality $x_{1,1} = 0$ induces cells:

$$\gamma_1 := D^{+,+,+,\dots,+}_{2,1,1,\dots,1}(1,2,3,\dots,k), \qquad \gamma_2 := D^{-,+,+,\dots,+}_{2,1,1,\dots,1}(1,2,3,\dots,k)$$

that are related, as sets, via $\gamma_2 = \varepsilon_1 \cdot \gamma_1$. Both cells γ_1 and γ_2 belong to the linear subspace

$$V_1 = \{(x_1,\dots,x_k) \in \mathbb{R}^{(d+1)\times k} : x_{1,1} = 0\}.$$

(B) The equality $x_{1,r-1} = x_{1,r}$ for $2 \le r \le k$ gives cells:

$$\gamma_{2r-1} := D^{+,+,+,\dots,+}_{1,\dots,1,2,1,\dots,1}(1,\dots,r-1,r,r+1,\dots,k),$$

$$\gamma_{2r} := D^{+,+,+,\dots,+}_{1,\dots,1,2,1,\dots,1}(1,\dots,r,r-1,r+1,\dots,k)$$

satisfying $\gamma_{2r} = \tau_{r-1,r}\cdot\gamma_{2r-1}$. In these cells the index 2 in the subscript $1,\dots,1,2,1,\dots,1$ appears at the position r. These cells belong to the linear subspace

$$V_r = \{(x_1,\dots,x_k) \in \mathbb{R}^{(d+1)\times k} : x_{1,r-1} = x_{1,r}\}.$$

Let e_θ denote a generator in $C_{N_2+1}(X_{d,k}, X^{>1}_{d,k})$ that corresponds to the cell θ. Furthermore let $e_{\gamma_1},\dots,e_{\gamma_{2k}}$ denote generators in $C_{N_2}(X_{d,k}, X^{>1}_{d,k})$ related to the cells $\gamma_1,\dots,\gamma_{2k}$.

The boundary of the cell θ is contained in the union of the linear subspaces $V_1,\dots,V_k$. Therefore we can orient the cells $\gamma_{2i-1},\gamma_{2i}$ consistently with the orientation of V_i, $1 \le i \le k$, that is given in such a way that

$$\partial e_\theta = (e_{\gamma_1} + e_{\gamma_2}) + (e_{\gamma_3} + e_{\gamma_4}) + \dots + (e_{\gamma_{2k-1}} + e_{\gamma_{2k}}).$$

Consequently,

$$\partial e_\theta = (1 + (-1)^d \varepsilon_1)\cdot e_{\gamma_1} + \sum_{i=2}^{k} (1 + (-1)^d \tau_{i-1,i})\cdot e_{\gamma_{2i-1}}. \tag{11}$$

(2) Let $\ell = 1$. Then $\theta := D^{+,+,+,\dots,+}_{2,1,1,\dots,1}(1,2,3,\dots,k)$. Now the cells in the boundary of θ are obtained by introducing extra equalities:

$$x_{2,1} = 0\,, \quad (0 =)\, x_{1,1} = x_{1,2}\,, \quad \dots \quad x_{1,k-1} = x_{1,k}.$$

Each of these equalities, except for the second one, will give two cells of dimension N_2, which yields $2(k-1)$ cells in total, in the boundary of θ. The equality $x_{1,1} = x_{1,2}$ will give additional four cells in the boundary of θ.

(A) The equality $x_{2,1} = 0$ induces cells:

$$\gamma_1 := D^{+,+,+,\dots,+}_{3,1,1,\dots,1}(1,2,3,\dots,k), \qquad \gamma_2 := D^{-,+,+,\dots,+}_{3,1,1,\dots,1}(1,2,3,\dots,k)$$

that are related, as sets, via $\gamma_2 = \varepsilon_1 \cdot \gamma_1$. Notice that both cells γ_1 and γ_2 belong to the linear subspace

$$V_1 = \{(x_1,\dots,x_k) \in \mathbb{R}^{(d+1)\times k} : x_{1,1} = x_{2,1} = 0\}.$$

(B) The equality $x_{1,1} = x_{1,2}$ yields the cells

$$\gamma_3 := D^{+,+,+,\dots,+}_{2,2,1,\dots,1}(1,2,3,\dots,k), \quad \gamma_{31} := D^{+,-,+,\dots,+}_{2,2,1,\dots,1}(1,2,3,\dots,k),$$

$$\gamma_{32} := D^{+,+,+,\dots,+}_{2,2,1,\dots,1}(2,1,3,\dots,k), \quad \gamma_{33} := D^{-,+,+,\dots,+}_{2,2,1,\dots,1}(2,1,3,\dots,k).$$

They satisfy set identities $\gamma_{31} = \varepsilon_2 \cdot \gamma_3$, $\gamma_{32} = \tau_{1,2} \cdot \gamma_3$, and $\gamma_{33} = \varepsilon_1 \tau_{1,2} \cdot \gamma_3$. All four cells belong to the linear subspace

$$V_2 = \{(x_1,\dots,x_k) \in \mathbb{R}^{(d+1)\times k} : 0 = x_{1,1} = x_{1,2}\}.$$

(C) The equality $x_{1,r-1} = x_{1,r}$ for $3 \le r \le k$ gives cells:

$$\gamma_{2r-1} := D^{+,+,+,\dots,+}_{2,\dots,1,2,1,\dots,1}(1,\dots,r-1,r,r+1,\dots,k),$$

$$\gamma_{2r} := D^{+,+,+,\dots,+}_{2,\dots,1,2,1,\dots,1}(1,\dots,r,r-1,r+1,\dots,k)$$

satisfying $\gamma_{2r} = \tau_{r-1,r} \cdot \gamma_{2r-1}$. In these cells the second index 2 in the subscript $2,\dots,1,2,1,\dots,1$ appears at the position r. These cells belong to the linear subspace

$$V_r = \{(x_1,\dots,x_k) \in \mathbb{R}^{(d+1)\times k} : x_{1,1} = 0,\ x_{1,r-1} = x_{1,r}\}.$$

Again e_θ denotes a generator in $C_{N_2+1}(X_{d,k}, X^{>1}_{d,k})$ corresponding to θ. Let $e_{\gamma_1}, e_{\gamma_2}, e_{\gamma_3}, e_{\gamma_{31}}, e_{\gamma_{32}}, e_{\gamma_{33}}, e_{\gamma_4}\dots, e_{\gamma_{2k}}$ denote generators in $C_{N_2}(X_{d,k}, X^{>1}_{d,k})$ related to the cells $\gamma_1, \gamma_2, \gamma_3, \gamma_{31}, \gamma_{32}, \gamma_{33}, \dots, \gamma_{2k}$.
The boundary of the cell θ, as before, is contained in the union of the linear subspaces $V_1, \dots, V_k$. Therefore we can orient cells consistently with the orientation of V_i, $1 \le i \le k$, that is given in such a way that

$$\partial e_\theta = (e_{\gamma_1} + e_{\gamma_2}) + (e_{\gamma_3} + e_{\gamma_{31}} + e_{\gamma_{32}} + e_{\gamma_{33}}) + \cdots + (e_{\gamma_{2k-1}} + e_{\gamma_{2k}}).$$

Consequently,

$$\begin{aligned} \partial e_\theta \quad = \quad & (1 + (-1)^{d-1}\varepsilon_1) \cdot e_{\gamma_1} + \\ & (1 + (-1)^d \varepsilon_2 + (-1)^d \tau_{1,2} + (-1)^{d+d} \varepsilon_1 \tau_{1,2}) \cdot e_{\gamma_3} + \\ & \sum_{i=3}^{k} (1 + (-1)^d \tau_{i-1,i}) \cdot e_{\gamma_{2i-1}}. \end{aligned} \tag{12}$$

(3) Let $2 \leq \ell \leq d-1$. Then $\theta := D^{+,+,+,\dots,+}_{\ell+1,1,1,\dots,1}(1,2,3,\dots,k)$. The cells in the boundary of θ are now obtained by introducing following equalities:

$$x_{\ell+1,1} = 0\,, \quad (0=)x_{1,1} = x_{1,2}\,, \quad \dots \quad x_{1,k-1} = x_{1,k}.$$

Each of them will give two cells of dimension N_2 in the boundary of θ, all together $2k$.

(A) The equality $x_{\ell+1,1} = 0$ induces cells:

$$\gamma_1 := D^{+,+,+,\dots,+}_{\ell+2,1,1,\dots,1}(1,2,3,\dots,k), \qquad \gamma_2 := D^{-,+,+,\dots,+}_{\ell+2,1,1,\dots,1}(1,2,3,\dots,k)$$

that are related, as sets, via $\gamma_2 = \varepsilon_1 \cdot \gamma_1$. Both cells γ_1 and γ_2 belong to the linear subspace

$$V_1 = \{(x_1,\dots,x_k) \in \mathbb{R}^{(d+1)\times k} : x_{1,1} = \dots = x_{\ell+1,1} = 0\}.$$

(B) The equality $(0=)x_{1,1} = x_{1,2}$ gives the cells

$$\gamma_3 := D^{+,+,+,\dots,+}_{\ell+1,2,1,\dots,1}(1,2,3,\dots,k), \qquad \gamma_4 := D^{+,-,+,\dots,+}_{\ell+1,2,1,\dots,1}(1,2,3,\dots,k)$$

that satisfy $\gamma_4 = \varepsilon_2 \cdot \gamma_3$. Both cells belong to the linear subspace

$$V_2 = \{(x_1,\dots,x_k) \in \mathbb{R}^{(d+1)\times k} : x_{1,1} = \dots = x_{\ell,1} = 0,\, x_{1,1} = x_{1,2}\}.$$

(C) The equality $x_{1,r-1} = x_{1,r}$ for $3 \leq r \leq k$ gives cells:

$$\gamma_{2r-1} := D^{+,+,+,\dots,+}_{\ell+1,\dots,1,2,1,\dots,1}(1,\dots,r-1,r,r+1,\dots,k),$$
$$\gamma_{2r} := D^{+,+,+,\dots,+}_{\ell+1,\dots,1,2,1,\dots,1}(1,\dots,r,r-1,r+1,\dots,k)$$

satisfying $\gamma_{2r} = \tau_{r-1,r} \cdot \gamma_{2r-1}$. In these cells the index 2 in the subscript $\ell+1,\dots,1,2,1,\dots,1$ appears at the position r. These cells belong to the linear subspace

$$V_r = \{(x_1,\dots,x_k) \in \mathbb{R}^{(d+1)\times k} : x_{1,1} = \dots = x_{\ell,1} = 0,\, x_{1,r-1} = x_{1,r}\}.$$

Again e_θ denotes a generator in $C_{N_2+1}(X_{d,k}, X^{>1}_{d,k})$ that corresponds to the cell θ. Furthermore $e_{\gamma_1},\dots,e_{\gamma_{2k}}$ denote generators in $C_{N_2}(X_{d,k}, X^{>1}_{d,k})$ related to the cells $\gamma_1,\dots,\gamma_{2k}$.
As before, the boundary of the cell θ is contained in the union of the linear subspaces $V_1,\dots,V_k$. Thus we can orient cells $\gamma_{2i-1},\gamma_{2i}$ consistently with the orientation of V_i, $1 \leq i \leq k$, that is given in such a way that

$$\partial e_\theta = (e_{\gamma_1} + e_{\gamma_2}) + (e_{\gamma_3} + e_{\gamma_4}) + \dots + (e_{\gamma_{2k-1}} + e_{\gamma_{2k}}).$$

Hence,

$$\partial e_\theta = (1+(-1)^{d-\ell}\varepsilon_1)\cdot e_{\gamma_1} + (1+(-1)^d\varepsilon_2)\cdot e_{\gamma_3} + \sum_{i=3}^{k}(1+(-1)^d\tau_{i-1,i})\cdot e_{\gamma_{2i-1}}. \tag{13}$$

The relations (11), (12) and (13) that we have now derived will be essential in the proofs of Theorems 1.4 and 1.5.

3.4 The arrangements parametrized by a cell

In this section we describe all arrangements of k hyperplanes parametrized by the cell

$$\theta := D^{+,+,+,\dots,+}_{\ell+1,1,1,\dots,1}(1,2,3,\dots,k),$$

where $1 \leq \ell \leq d-1$. This description will be one of the key ingredients in Section 4 when the obstruction cocycle is evaluated on the cell θ.
Recall that the cell θ is defined as the intersection of the sphere $S(\mathbb{R}^{(d+1)\times k})$ and the cone given by the inequalities:

$$0 <_{\ell+1} x_1 <_1 x_2 <_1 \cdots <_1 x_k.$$

Consider the binomial coefficient moment curve $\hat{\gamma}\colon \mathbb{R} \longrightarrow \mathbb{R}^d$ defined by

$$\hat{\gamma}(t) = (t, \tbinom{t}{2}, \tbinom{t}{3}, \dots, \tbinom{t}{d})^t. \tag{14}$$

After embedding $\mathbb{R}^d \longrightarrow \mathbb{R}^{d+1}, (\xi_1,\dots,\xi_d)^t \longmapsto (1,\xi_1,\dots,\xi_d)^t$ it corresponds to the curve $\gamma\colon \mathbb{R} \longrightarrow \mathbb{R}^{d+1}$ given

$$\gamma(t) = (1, t, \tbinom{t}{2}, \tbinom{t}{3}, \dots, \tbinom{t}{d})^t.$$

Consider the following points on the moment curve γ:

$$q_1 := \gamma(0), \dots, q_{\ell+1} := \gamma(\ell). \tag{15}$$

Next, recall that each oriented affine hyperplane $\hat{H}$ in $\mathbb{R}^d$ (embedded in $\mathbb{R}^{d+1}$) determines the unique linear hyperplane H such that $\hat{H} = H \cap \mathbb{R}^d$, and almost vice versa. Now, the family of arrangements parametrized by the (open) cell θ is described as follows:

Lemma 3.13. *The cell $\theta = D^{+,+,+,\dots,+}_{\ell+1,1,1,\dots,1}(1,2,3,\dots,k)$ parametrizes all arrangements $\mathcal{H} = (H_1,\dots,H_k)$ of k linear hyperplanes in $\mathbb{R}^{d+1}$, where the order and orientation are fixed appropriately such that*

- $Q := \{q_1,\dots,q_\ell\} \subset H_1$,
- $q_{\ell+1} \notin H_1$,
- $q_1 \notin H_2, \dots, q_1 \notin H_k$, *and*
- $H_2,\dots,H_k$ *have unit normal vectors with different (positive) first coordinates, that is,* $|\{\langle x_2,q_1\rangle, \langle x_3,q_1\rangle, \dots, \langle x_k,q_1\rangle\}| = k-1$.

Here $x_i \in S(\mathbb{R}^{(d+1)\times k})$ is a unit normal vector of the hyperplane H_i, for $1 \leq i \leq k$.

Proof. Observe that $\{q_1,\dots,q_\ell\} \subset H_1$ holds if and only if $\langle x_1,q_1\rangle = \langle x_1,q_2\rangle = \cdots = \langle x_1,q_\ell\rangle = 0$ if and only if $x_{1,1} = x_{2,1} = \cdots = x_{\ell,1} = 0$: This is true since we have the binomial moment curve, so $q_i = \gamma(i-1)$ has only the first i coordinates non-zero.
Furthermore, $q_{\ell+1} \notin H_1$ holds if and only if $x_{\ell+1,1} \neq 0$; choosing an appropriate orientation for H_1 we can assume that $x_{\ell+1,1} > 0$.

The third condition is equivalent to $0 \notin \{\langle x_2, q_1\rangle, \langle x_3, q_1\rangle, \dots, \langle x_k, q_1\rangle\}$, that is, $x_{1,2}, x_{1,3}, \dots, x_{1,k} \neq 0$. Choosing orientations of $H_2, \dots, H_k$ suitably this yields $x_{1,2}, x_{1,3}, \dots, x_{1,k} > 0$.
Since the values $x_{1,2} = \langle x_2, q_1\rangle$, $x_{1,3} = \langle x_3, q_1\rangle$, $\dots$, $x_{1,k} = \langle x_k, q_1\rangle$ are positive and distinct, we get $0 < x_{1,2} < x_{1,3} < \cdots < x_{1,k}$ by choosing the right order on $H_2, \dots, H_k$. ☐

4 Proofs

4.1 Proof of Theorem 1.3

Let $d \geq 1$, $j \geq 1$, $\ell \geq 0$ and $k \geq 2$ be integers with the property that $dk = j(2^k - 1) + \ell$ for $0 \leq \ell \leq d-1$.
Consider a collection of j ordered disjoint intervals $\mathcal{M} = (I_1, \dots, I_j)$ along the moment curve γ. Let $Q = \{q_1, \dots, q_\ell\} \subset \gamma$ be a set of ℓ predetermined points that lie to the left of the interval I_1. We prove Theorem 1.3 in two steps.

Lemma 4.1. *Let A be an ℓ-equiparting matrix, that is, a binary matrix of size $k \times j2^k$ with one row of transition count $d - \ell$ and all other rows of transition count d such that $A = (A_1, \dots, A_j)$ for Gray codes $A_1, \dots, A_j$ with the property that the last column of A_i is equal to the first column of A_{i+1} for $1 \leq i < j$. Then A determines an arrangement $\mathcal{H}$ of k affine hyperplanes that equipart $\mathcal{M} = (I_1, \dots, I_j)$ and one of the hyperplanes passes through each point in Q.*

Proof. Without loss of generality we assume that the first row of the matrix A has transition count $d - \ell$ while rows 2 through k have transition count d. For a row a_s of the matrix A, denote by t_s its transition count, $1 \leq s \leq k$.
Place $j(2^k + 1)$ ordered points $q_{\ell+1}, \dots, q_{\ell+j(2^k+1)}$ on γ, such that

$$I_i = [q_{\ell+(i-1)2^k+i},\ q_{\ell+i2^k+i}]$$

and each sequence of 2^k+1 points divides I_i into 2^k subintervals of equal length. *Ordered* refers to the property that $q_r = \gamma(t_r)$ if $t_1 < t_2 < \cdots < t_{j(2^k+1)}$.
We now define the hyperplanes in $\mathcal{H}$ by specifying which of the points they pass through and then choosing their orientations. Force the affine hyperplane H_1 to pass through all of the points in Q. For $s = 1, \dots, i$, the affine hyperplane H_s passes through $x_{\ell+r+i}$ if there is a bit change in row a_s from entry r to entry $r+1$ for $(i-1)2^k < r \leq i2^k$. Orient H_s such that the subinterval $[q_r, q_{r+1}]$ is on the positive side of H_s if it corresponds to a 0-entry in a_s. Since each $A_1, \dots, A_j$ is a Gray code, the arrangement $\mathcal{H}$ is indeed an equipartition. ☐

Lemma 4.2. *Every arrangement of k affine hyperplanes $\mathcal{H}$ that equiparts $\mathcal{M} = (I_1, \dots, I_j)$ and where one of the hyperplanes passes through each point of Q induces a unique binary matrix A as in Lemma 4.1.*

Proof. Since $dk = j(2^k - 1) + \ell$ and $0 \leq \ell \leq d - 1$, the hyperplanes in $\mathcal{H}$ must pass through the points $q_{\ell+(i-1)2^k+i+1}, \dots, q_{\ell+i2^k+i-1}$ of the intervals I_i

for $i \in \{1, \ldots, j\}$. Recording the position of the subintervals $[q_{\ell+r}, q_{\ell+r+1}]$, for $r \neq i2^k + i$, with respect to each hyperplane leads to a matrix as in described in Lemma 4.1. $\square$

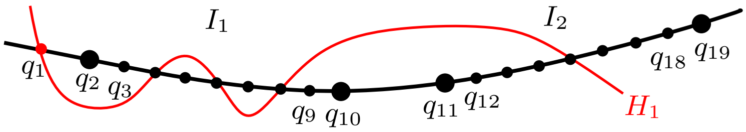


Figure 2: Illustration of one step in the proof of Lemma 4.1. Here H_1 is an affine hyperplane bisecting two intervals I_1 and I_2 on the 5-dimensional moment curve.

Thus the number of non-equivalent ℓ-equiparting matrices is the same as the number of arrangements of k affine hyperplanes $\mathcal{H}$ that equipart the collection of j disjoint intervals on the moment curve in $\mathbb{R}^d$, up to renumbering and orientation change of hyperplanes in $\mathcal{H}$, when one of the hyperplanes is forced to pass through ℓ prescribed points on the moment curve lying to the left of the intervals. This concludes the proof of Theorem 1.3. $\square$

4.2 Proof of Theorem 1.4

Let $j \geq 1$ and $k \geq 3$ be integers with $d = \lceil \frac{2^k-1}{k} j \rceil$ and $\ell = dk - (2^k - 1)j$. In addition, assume that the number of non-equivalent ℓ-equiparting matrices of size $k \times j2^k$ is odd.

In order to prove that $\Delta(j,k) \leq d$ it suffices by Theorem 2.3 to prove that there is no $\mathfrak{S}_k^{\pm}$-equivariant map

$$X_{d,k} \longrightarrow S(W_k \oplus U_k^{\oplus j}),$$

whose restriction to $X_{d,k}^{>1}$ is $\mathfrak{S}_k^{\pm}$-homotopic to $\nu \circ \psi_{\mathcal{M}}|_{X_{d,k}^{>1}}$ for $\mathcal{M} = (I_1, \ldots, I_j)$. Following Section 2.6 we verify that the cohomology class

$$[\mathfrak{o}(g)] \in \mathcal{H}^{N_2+1}_{\mathfrak{S}_k^{\pm}}\big(X_{d,k}, X_{d,k}^{>1}; \pi_{N_2}(S(W_k \oplus U_k^{\oplus j}))\big),$$

does not vanish, where $g = \nu \circ \psi_{\mathcal{M}}|_{X^{(N_2)}}$.

Consider the cell $\theta := D^{+,+,+,\ldots,+}_{\ell+1,1,1,\ldots,1}(1,2,3,\ldots,k)$ of dimension $(d+1)k-1-\ell = N_2+1$ in $X_{d,k}$, as in Example 3.12. Let e_θ denote the corresponding basis element of the cell θ in the cellular chain group $C_{N_2+1}(X_{d,k}, X_{d,k}^{>1})$, and let h_θ be the attaching map of θ. This cell is cut out from the unit sphere $S(\mathbb{R}^{(d+1)\times k})$ by the following inequalities:

$$0 <_{\ell+1} x_1 <_1 x_2 <_1 \cdots <_1 x_k.$$

In particular, this means that the first ℓ coordinates of x_1 are zero, i.e., $x_{1,1} = x_{2,1} = x_{3,1} = \cdots = x_{\ell,1} = 0$, and $x_{\ell+1,1} > 0$.

Let us fix ℓ points $Q = \{q_1, \dots, q_\ell\}$ on the moment curve (14) in $\mathbb{R}^{d+1}$ as it was done in (15): $q_1 := \gamma(0), \dots, q_\ell := \gamma(\ell - 1)$. Then, by Lemma 3.13, the relative interior of $D^{+,+,+,\dots,+}_{\ell+1,1,1,\dots,1}(1, 2, 3, \dots, k)$ parametrizes the arrangements $\mathcal{H} = (H_1, \dots, H_k)$ for which orientations and order of the hyperplanes are fixed with H_1 containing all the points from Q. According to the formula (5) we have that

$$\mathfrak{o}(g)(e_\theta) = [\nu \circ \psi_{\mathcal{M}} \circ h_\theta|_{\partial\theta}] = \sum \deg(\nu \circ \psi_{\mathcal{M}}|_{X^{(N_2+1)}} \circ h_\theta|_{S_i}) \cdot \zeta,$$

where as before $\zeta \in \pi_{N_2}(S(W_k \oplus U_k^{\oplus j})) \cong \mathbb{Z}$ is a generator, and the sum ranges over all arrangements of k hyperplanes in relint θ that equipart $\mathcal{M}$. Here, as before, S_i denotes a small N_2-sphere around a root of the function $\psi_{\mathcal{M}}|_{X^{(N_2+1)}} \circ h_\theta$, i.e., the point that parametrizes an arrangements of k hyperplanes in relint θ that equipart $\mathcal{M}$.

Now, the local degrees of the function $\nu \circ \psi_{\mathcal{M}}|_{X^{(N_2+1)}} \circ h_\theta$ are ± 1. Indeed, in a small neighborhood $U \subseteq$ relint θ around any root the test map $\psi_{\mathcal{M}}$ is a continuous bijection. Thus $\psi_{\mathcal{M}}|_{\partial U}$ is a continuous bijection into some N_2-sphere around the origin in $W_k \oplus U_k^{\oplus j}$ and by compactness of ∂U is a homeomorphism. Consequently,

$$\mathfrak{o}(g)(e_\theta) = \sum \deg(\nu \circ \psi_{\mathcal{M}}|_{X^{(N_2+1)}} \circ h_\theta|_{S_i}) \cdot \zeta = \big(\sum \pm 1\big) \cdot \zeta = a \cdot \zeta, \quad (16)$$

where the sum ranges over all arrangements of k hyperplanes in relint θ that equipart $\mathcal{M}$. According to Theorem 1.3 the number of (± 1)'s in the sum (16) is equal to the number of non-equivalent ℓ-equiparting matrices of size $k \times j2^k$. By our assumption this number is odd and consequently $a \in \mathbb{Z}$ is an odd integer. We obtained that

$$\mathfrak{o}(g)(e_\theta) = a \cdot \zeta, \quad (17)$$

where $a \in \mathbb{Z}$ is an odd integer.

Remark 4.3. It is important to point out that the calculations and formulas up to this point also hold for $k = 2$. The assumption $k \geq 3$ affects the $\mathfrak{S}_k^{\pm} = (\mathbb{Z}/2)^k \rtimes \mathfrak{S}_k$ module structure on $\pi_{N_2}(S(W_k \oplus U_k^{\oplus j})) \cong \mathbb{Z}$. For $k \geq 2$ every generator ε_i of the subgroup $(\mathbb{Z}/2)^k$ acts trivially, while each transposition $\tau_{i,t}$, a generator of the subgroup $\mathfrak{S}_k$, acts as multiplication by -1 in the case $k \geq 3$, and as multiplication by $(-1)^{j+1}$ in the case $k = 2$.

Finally, we prove that $[\mathfrak{o}(g)]$ does not vanish and conclude the proof. This will be achieved by proving that the cocycle $\mathfrak{o}(g)$ is not a coboundary.

Let us assume to the contrary that $\mathfrak{o}(g)$ is a coboundary. Thus there exists a cochain

$$\mathfrak{h} \in \mathcal{C}^{N_2}_{\mathfrak{S}_k^{\pm}}\big(X_{d,k}, X_{d,k}^{>1}\,;\, \pi_{N_2}(S(W_k \oplus U_k^{\oplus j}))\big)$$

such that $\mathfrak{o}(g) = \delta\mathfrak{h}$, where δ denotes the coboundary operator. In the case when

(1) $\ell = 0$ the relation (11) implies that

$$\begin{aligned} a\cdot\zeta &= \mathfrak{o}(g)(e_\theta) = \delta\mathfrak{h}(e_\theta) = \mathfrak{h}(\partial e_\theta) \\ &= (1+(-1)^d\varepsilon_1)\cdot\mathfrak{h}(e_{\gamma_1}) + \sum_{i=2}^{k}(1+(-1)^d\tau_{i-1,i})\cdot\mathfrak{h}(e_{\gamma_{2i-1}}) \\ &= (1+(-1)^d)\cdot\mathfrak{h}(e_{\gamma_1}) + \sum_{i=2}^{k}(1+(-1)^{d+1})\cdot\mathfrak{h}(e_{\gamma_{2i-1}}) \\ &= 2b\cdot\zeta, \end{aligned}$$

for some integer b. Since a is an odd integer this is not possible, and therefore $\mathfrak{o}(g)$ is not a coboundary.

(2) $\ell = 1$ the relation (12) implies that

$$\begin{aligned} a\cdot\zeta &= \mathfrak{o}(g)(e_\theta) = \delta\mathfrak{h}(e_\theta) = \mathfrak{h}(\partial e_\theta) \\ &= (1+(-1)^{d-1}\varepsilon_1)\cdot\mathfrak{h}(e_{\gamma_1}) + \\ &\quad (1+(-1)^d\varepsilon_2+(-1)^d\tau_{1,2}+(-1)^{d+d}\varepsilon_1\tau_{1,2})\cdot\mathfrak{h}(e_{\gamma_3}) + \\ &\quad \sum_{i=3}^{k}(1+(-1)^d\tau_{i-1,i})\cdot\mathfrak{h}(e_{\gamma_{2i-1}}) \\ &= (1+(-1)^{d-1})\cdot\mathfrak{h}(e_{\gamma_1}) + (1+(-1)^d+(-1)^{d+1}-1)\cdot\mathfrak{h}(e_{\gamma_3}) + \\ &\quad \sum_{i=3}^{k}(1+(-1)^{d+1})\cdot\mathfrak{h}(e_{\gamma_{2i-1}}) \\ &= (1+(-1)^{d-1})\cdot\mathfrak{h}(e_{\gamma_1}) + \sum_{i=3}^{k}(1+(-1)^{d+1})\cdot\mathfrak{h}(e_{\gamma_{2i-1}}) \\ &= 2b\cdot\zeta, \end{aligned}$$

for $b\in\mathbb{Z}$. Again we reached a contradiction, so $\mathfrak{o}(g)$ is not a coboundary.

(3) $2 \le \ell \le d-1$ the relation (13) implies that

$$\begin{aligned} a\cdot\zeta &= \mathfrak{o}(g)(e_\theta) = \delta\mathfrak{h}(e_\theta) = \mathfrak{h}(\partial e_\theta) \\ &= (1+(-1)^{d-\ell}\varepsilon_1)\cdot\mathfrak{h}(e_{\gamma_1}) + (1+(-1)^d\varepsilon_2)\cdot\mathfrak{h}(e_{\gamma_3}) + \\ &\quad \sum_{i=3}^{k}(1+(-1)^d\tau_{i-1,i})\cdot\mathfrak{h}(e_{\gamma_{2i-1}}) \\ &= (1+(-1)^{d-\ell})\cdot\mathfrak{h}(e_{\gamma_1}) + (1+(-1)^d)\cdot\mathfrak{h}(e_{\gamma_3}) + \\ &\quad \sum_{i=3}^{k}(1+(-1)^{d+1})\cdot\mathfrak{h}(e_{\gamma_{2i-1}}) \\ &= 2b\cdot\zeta, \end{aligned}$$

for an integer b. Since a is an odd integer this is not possible. Again, $\mathfrak{o}(g)$ is not a coboundary. ∎

4.3 Proof of Theorem 1.5

Let $j \geq 1$ be an integer with $d = \lceil \frac{3}{2}j \rceil$ and $\ell = 2d - 3j \leq 1$.
The proof of this theorem is done in the footsteps of the proof of Theorem 1.4. In all three cases we rely on Theorem 2.3 and prove

- the non-existence of $\mathfrak{S}_2^{\pm}$-equivariant map $X_{d,2} \longrightarrow S(W_2 \oplus U_2^{\oplus j})$ whose restriction to $X_{d,2}^{>1}$ is $\mathfrak{S}_2^{\pm}$-homotopic to $\nu \circ \psi_{\mathcal{M}}|_{X_{d,2}^{>1}}$ for $\mathcal{M} = (I_1, \ldots, I_j)$
- by evaluating the obstruction cocycle $\mathfrak{o}(g)$ for $g = \nu \circ \psi_{\mathcal{M}}|_{X^{(N_2)}}$ on cells $D_{1,1}^{+,+}(1,2)$ or $D_{2,1}^{+,+}(1,2)$, depending on ℓ being 0 or 1, using Theorem 1.3; and then
- prove that the cocycle $\mathfrak{o}(g)$ cannot be a coboundary, utilizing boundary formulas from Example 3.12.

4.3.1 2-bit Gray codes

In order to evaluate the obstruction cocycle $\mathfrak{o}(g)$ on the relevant cells in the case $k = 2$ we need to understand (2×4)-Gray codes. These correspond to equipartitions of an interval I on the moment curve into four equal orthants by intersecting with two hyperplanes H_1 and H_2 in altogether three points of the interval. There are two such configurations: either H_1 cuts through the midpoint of I and H_2 separates both halves of I into equal pieces by two additional intersections, or the roles of H_1 and H_2 are reversed. In terms of Gray codes we can express this as follows.

Lemma 4.4. *There are two different 2-bit Gray codes that start with the zero column (or any other fixed binary vector of length 2):*

$$\begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

Proof. The second column of the Gray code determines the rest of the code, and there are only two choices for a bit flip. $\square$

This means that in the case $k = 2$ an ℓ-equiparting matrix A has a more compact representation: it is determined by the first column – a binary vector of length 2 – and j additional bits, one for each A_i, encoding whether the first bit flip in A_i is in the first or second row. These j bits *cannot be chosen independently* since there are restrictions imposed by the transition count.

Lemma 4.5. *Let $j \geq 1$ be an integer with $d = \lceil \frac{3}{2}j \rceil$ and $\ell = 2d - 3j \leq 1$.*

(1) *If $\ell = 0$, then the number of non-equivalent 0-equiparting matrices is equal to*
$$\tfrac{1}{2}\binom{j}{\frac{j}{2}}.$$

(2) *If $\ell = 1$, then the number of non-equivalent 1-equiparting matrices is equal to*
$$\binom{j}{\frac{j+1}{2}}.$$

Proof. We count the number of non-equivalent ℓ-equiparting matrices of the form $A = (A_1, \dots, A_j)$ where A_i is a 2-bit Gray code. A (2×4)-Gray code with the first bit flip in the first row has in total two bit flips in the first row and one bit flip in the second row.

(1) Let $\ell = 0$. Then $2d = 3j$ and consequently j has to be even. The matrix A must have transition count d in each row. Thus, half of the A_i's have the first bit flip in the first row. Consequently, 0-equiparting matrices A with a fixed first column are in bijection with $\frac{j}{2}$-element subsets of a set with j elements. By inverting the bits in each row we can fix the first column of A to be the zero vector. Additionally, we are allowed to interchange the rows. Up to this equivalence there are $\frac{1}{2}\binom{j}{j/2}$ such matrices.

(2) Let $\ell = 1$. Then $2d = 3j + 1$ and so j is odd. The matrix A must have transition count d in one row while transition count $d - 1$ in the remaining row. Without loss of generality we can assume that A have transition count d in the first row. Assume that r of the A_i's have the first bit flip in the first row. Consequently, $j - r$ of the A_i's have the first bit flip in the second row. Now the transition count of the first row is $2r + j - r$ while the transition count of the second row is $r + 2(j - r)$. The system of equations $2r + j - r = d$, $r + 2(j - r) = d - 1$ yields that $r = \frac{j+1}{2}$. Therefore, up to equivalence, there are $\binom{j}{r}$ such matrices. $\square$

4.3.2 The case $\ell = 0 \Leftrightarrow 2d = 3j$

Let $\theta := D^{+,+}_{1,1}(1,2)$, and let e_θ denote the related basis element of the cell θ in the top cellular chain group $C_{2d+1}(X_{d,2}, X^{>1}_{d,2})$ which, in this case, is equivariantly generated by θ. According to equation (16), which also holds for $k = 2$ as explained in Remark 4.3,

$$\mathfrak{o}(g)(e_\theta) = \big(\sum \pm 1\big) \cdot \zeta = a \cdot \zeta, \tag{18}$$

where $\zeta \in \pi_{2d+1}(S(W_2 \oplus U_2^{\oplus j})) \cong \mathbb{Z}$ is a generator, and the sum ranges over all arrangements of two hyperplanes in $\operatorname{relint}\theta$ that equipart $\mathcal{M}$. Since θ parametrizes all arrangements $\mathcal{H} = (H_1, H_2)$ where orientations and order of hyperplanes are fixed, the sum in (18) ranges over all arrangements of two hyperplanes that equipart $\mathcal{M}$ where orientation and order of hyperplanes are fixed. Therefore, by Theorem 1.3, the number of (± 1)'s in the sum of (18) is equal to the number of non-equivalent 0-equiparting matrices of size $2 \times 4j$. Now, Lemma 4.5 implies that the number of (± 1)'s in the sum of (18) is $\frac{1}{2}\binom{j}{j/2}$. Consequently, integer a is odd if and only if $\frac{1}{2}\binom{j}{j/2}$ is odd.

Assume that the cocycle $\mathfrak{o}(g)$ is a coboundary. Hence, there exists a cochain

$$\mathfrak{h} \in \mathcal{C}^{2d}_{\mathfrak{S}_2^{\pm}}\big(X_{d,2}, X^{>1}_{d,2}\,;\, \pi_{2d}(S(W_2 \oplus U_2^{\oplus j}))\big)$$

with the property that $\mathfrak{o}(g) = \delta\mathfrak{h}$. The relation (11) for $k = 2$ transforms into

$$\partial e_\theta = (1 + (-1)^d \varepsilon_1) \cdot e_{\gamma_1} + (1 + (-1)^d \tau_{1,2}) \cdot e_{\gamma_3}.$$

Thus we have that

$$\begin{aligned} a\cdot\zeta &= \mathfrak{o}(g)(e_\theta)=\delta\mathfrak{h}(e_\theta)=\mathfrak{h}(\partial e_\theta)\\ &= (1+(-1)^d\varepsilon_1)\cdot\mathfrak{h}(e_{\gamma_1})+(1+(-1)^d\tau_{1,2})\cdot\mathfrak{h}(e_{\gamma_3})\\ &= (1+(-1)^d)\cdot\mathfrak{h}(e_{\gamma_1})+(1+(-1)^{d+j+1})\cdot\mathfrak{h}(e_{\gamma_3})\\ &= 2b\cdot\zeta. \end{aligned}$$

Consequently, $\mathfrak{o}(g)$ is not a coboundary if and only if a is odd if and only if $\frac{1}{2}\binom{j}{j/2}$ is odd. Having in mind the Kummer criterion stated below we conclude that: A $\mathfrak{S}_2^{\pm}$-equivariant map $X_{d,2}\longrightarrow S(W_2\oplus U_2^{\oplus j})$ whose restriction to $X_{d,2}^{>1}$ is $\mathfrak{S}_2^{\pm}$-homotopic to $\nu\circ\psi_{\mathcal{M}}|_{X_{d,2}^{>1}}$ does not exists *if and only is* $\mathfrak{o}(g)$ is not a coboundary *if and only if* a is an odd integer *if and only if* $\frac{1}{2}\binom{j}{j/2}$ is odd *if and only if* $j=2^t$ for $t\geq 1$.

Lemma 4.6 (Kummer [12]). *Let $n\geq m\geq 0$ be integers and let p be a prime. The maximal integer k such that p^k divides $\binom{n}{m}$ is the number of carries when m and $n-m$ are added in base p.*

Thus we have proved the case (ii) of Theorem 1.5. Moreover, since the primary obstruction $\mathfrak{o}(g)$ is the only obstruction, we have proved that a $\mathfrak{S}_2^{\pm}$-equivariant map $X_{d,2}\longrightarrow S(W_2\oplus U_2^{\oplus j})$ whose restriction to $X_{d,2}^{>1}$ is $\mathfrak{S}_2^{\pm}$-homotopic to $\nu\circ\psi_{\mathcal{M}}|_{X_{d,2}^{>1}}$ exists if and only if j, an even integer, is not a power of 2.

4.3.3 The case $\ell=1\Leftrightarrow 2d=3j+1$

Let $\theta:=D_{2,1}^{+;+}(1,2)$, and again let e_θ denote the related basis element of the cell θ in the cellular chain group $C_{2d}(X_{d,2},X_{d,2}^{>1})$ which, in this case, is equivariantly generated by two cells $D_{2,1}^{+;+}(1,2)$ and $D_{1,2}^{+;+}(1,2)$. Again, the equation (16) implies that

$$\mathfrak{o}(g)(e_\theta)=\big(\sum\pm 1\big)\cdot\zeta=a\cdot\zeta, \tag{19}$$

where $\zeta\in\pi_{2d+1}(S(W_2\oplus U_2^{\oplus j}))\cong\mathbb{Z}$ is a generator, and the sum ranges over all arrangements of k hyperplanes in $\operatorname{relint}\theta$ that equipart $\mathcal{M}$. The cell θ parametrizes all arrangements $\mathcal{H}=(H_1,H_2)$ where H_1 passes through the given point on the moment curve and orientations and order of hyperplanes are fixed. Thus, the sum in (19) ranges over all arrangements of two hyperplanes that equipart $\mathcal{M}$ where H_1 passes through the given point on the moment curve with order and orientation of hyperplanes being fixed. Therefore, by Theorem 1.3, the number of (±1)'s in the sum of (19) is the same as the number of non-equivalent 1-equiparting matrices of size $2\times 4j$. Again, Lemma 4.5 implies that the number of (±1)'s in the sum of (19) is $\binom{j}{(j+1)/2}$. The integer a is odd if and only if $\binom{j}{(j+1)/2}$ is odd if and only if $j=2^t-1$ for $t\geq 1$.
Assume that the cocycle $\mathfrak{o}(g)$ is a coboundary. Then there exists a cochain

$$\mathfrak{h}\in\mathcal{C}^{2d-1}_{\mathfrak{S}_2^{\pm}}\big(X_{d,2},X_{d,2}^{>1}\,;\,\pi_{2d-1}(S(W_2\oplus U_2^{\oplus j}))\big)$$

with the property that $\mathfrak{o}(g) = \delta\mathfrak{h}$. Now, the relation (12) for $k = 2$ transforms into

$$\partial e_\theta = (1 + (-1)^{d-1}\varepsilon_1) \cdot e_{\gamma_1} + (1 + (-1)^d\varepsilon_2 + (-1)^d\tau_{1,2} + (-1)^{d+d}\varepsilon_1\tau_{1,2}) \cdot e_{\gamma_3}.$$

Thus, having in mind that j has to be odd, we have

$$\begin{aligned} a \cdot \zeta &= \mathfrak{o}(g)(e_\theta) = \delta\mathfrak{h}(e_\theta) = \mathfrak{h}(\partial e_\theta) \\ &= (1 + (-1)^{d-1}\varepsilon_1) \cdot \mathfrak{h}(e_{\gamma_1}) + \\ &\quad (1 + (-1)^d\varepsilon_2 + (-1)^d\tau_{1,2} + (-1)^{d+d}\varepsilon_1\tau_{1,2}) \cdot \mathfrak{h}(e_{\gamma_3}) \\ &= (1 + (-1)^{d-1}) \cdot \mathfrak{h}(e_{\gamma_1}) + (1 + (-1)^d + (-1)^{d+j+1} + (-1)^{j+1}) \cdot \mathfrak{h}(e_{\gamma_3}) \\ &= (1 + (-1)^{d-1}) \cdot \mathfrak{h}(e_{\gamma_1}) + (1 + (-1)^d + (-1)^d + 1) \cdot \mathfrak{h}(e_{\gamma_3}) \\ &= \begin{cases} 2\mathfrak{h}(e_{\gamma_1}), & d \text{ odd} \\ 4\mathfrak{h}(e_{\gamma_3}), & d \text{ even.} \end{cases} \end{aligned} \tag{20}$$

Now, we separately consider cases depending on parity of d and value of j.

(1) Let d be odd. Recall that a is odd if and only if $j = 2^t - 1$ for $t \geq 1$. Since $d = \frac{1}{2}(3j+1) = 3 \cdot 2^{t-1} - 1$ and d is odd we have that for $j = 2^t - 1$, with $t \geq 2$, the integer a is odd and consequently $\mathfrak{o}(g)$ is not a coboundary. Thus a $\mathfrak{S}_2^{\pm}$-equivariant map $X_{d,2} \longrightarrow S(W_2 \oplus U_2^{\oplus j})$ whose restriction to $X_{d,2}^{>1}$ is $\mathfrak{S}_2^{\pm}$-homotopic to $\nu \circ \psi_{\mathcal{M}}|_{X_{d,2}^{>1}}$ does not exists. We have proved the case (ii) of Theorem 1.5 for $t \geq 2$.

(2) Let $d = 2$ and $j = 1 = 2^1 - 1$. Then the integer a is again odd and consequently cannot be divisible by 4 implying again that $\mathfrak{o}(g)$ is not a coboundary. Therefore a $\mathfrak{S}_2^{\pm}$-equivariant map $X_{2,2} \longrightarrow S(W_2 \oplus U_2)$ whose restriction to $X_{2,2}^{>1}$ is $\mathfrak{S}_2^{\pm}$-homotopic to $\nu \circ \psi_{\mathcal{M}}|_{X_{2,2}^{>1}}$ does not exists. This concludes the proof of the case (ii) of Theorem 1.5.

(3) Let $d \geq 4$ be even. Now we determine the integer a by computing local degrees $\deg(\nu \circ \psi_{\mathcal{M}}|_{X^{(N_2+1)}} \circ h_\theta|_{S_i})$; see (16) and (19). We prove, almost identically as in [3, Proof of Lem. 5.6], that all local degrees equal, either 1 or -1.

That local degrees of $\nu \circ \psi_{\mathcal{M}}|_\theta$ are ± 1 is simple to see since in a small neighborhood U in $\operatorname{relint}\theta$ around any root $\lambda u + (1-\lambda)v$ the test map $\psi_{\mathcal{M}}|_\theta$ is a continuous bijection. Indeed, for any vector $w \in W_2 \oplus U_2^{\oplus j}$, with sufficiently small norm, there is exactly one $\lambda u' + (1-\lambda)v' \in U$ with $\psi_{\mathcal{M}}(\lambda u' + (1-\lambda)v') = w$. Thus $\psi_{\mathcal{M}}|_{\partial U}$ is a continuous bijection into some $3j$-sphere around the origin of $W_2 \oplus U_2^{\oplus j}$ and by compactness of ∂U is a homeomorphism.

Next we compute the signs of the local degrees. First we describe a neighborhood of every root of the test map $\psi_{\mathcal{M}}$ in $\operatorname{relint}\theta$. Let $\lambda u + (1-\lambda)v \in \operatorname{relint}\theta$ with $\psi_{\mathcal{M}}(\lambda u + (1-\lambda)v) = 0$. Consequently $\lambda = \frac{1}{2}$. Denote the intersections of the hyperplane H_u with the moment curve by $x_1, \ldots, x_d$ in the correct order along the moment curve. Similarly, let $y_1, \ldots, y_d$ be the intersections of H_v with the moment curve. In particular, x_1 is the point q_1 that determines the cell θ, see Lemma 3.13. Choose an $\epsilon > 0$ such that ϵ-balls around $x_2, \dots, x_d$ and around $y_1, \dots, y_d$ are pairwise disjoint with the property that these balls intersect the moment curve only in precisely one of the intervals $I_1, \dots, I_j$. Pairs of hyperplanes $(H_{u'}, H_{v'})$ with $\lambda u' + (1-\lambda)v' \in \operatorname{relint}\theta$ that still intersect the moment curve in the corresponding ϵ-balls parametrize a neighborhood of $\frac{1}{2}u + \frac{1}{2}v$. The local neighborhood consisting of pairs of hyperplanes with the same orientation still intersecting the moment curve in the corresponding ϵ-balls where the parameter λ is in some neighborhood of $\frac{1}{2}$. For sufficiently small $\epsilon > 0$ the neighborhood can be naturally parametrized by the product

$$(\tfrac{1}{2} - \epsilon, \tfrac{1}{2} + \epsilon) \times \prod_{i=2}^{2d} (-\epsilon, \epsilon),$$

where the first factor relates to λ, the next $d-1$ factors correspond to neighborhoods of the $x_2, \dots, x_d$ and the last d factors to ϵ-balls around $y_1, \dots, y_d$. A natural basis of the tangent space at $\frac{1}{2}u + \frac{1}{2}v$ is obtained via the push-forward of the canonical basis of $\mathbb{R}^{2d}$ as tangent space at $(\frac{1}{2}, 0, \dots, 0)^t$.
Consider the subspace $Z \subseteq \operatorname{relint}\theta$ that consists all points $\lambda u + (1-\lambda)v$ associated to the pairs of hyperplanes (H_u, H_v) such that both hyperplanes intersect the moment curve in d points. In the space Z the local degrees only depend on the orientations of the hyperplanes H_u and H_v, but these are fixed since $Z \subseteq \operatorname{relint}\theta$. Indeed, any two neighborhoods of distinct roots of the test map $\psi_{\mathcal{M}}$ can be mapped onto each other by a composition of coordinate charts since their domains coincide. This is a smooth map of degree 1: the Jacobian at the root is the identity map. Let $\frac{1}{2}u + \frac{1}{2}v$ and $\frac{1}{2}u' + \frac{1}{2}v'$ be roots in Z of the test map $\psi_{\mathcal{M}}$ and let Ψ be the change of coordinate chart described above. Then $\psi_{\mathcal{M}}$ and $\psi_{\mathcal{M}} \circ \Psi$ differ in a neighborhood of $\frac{1}{2}u + \frac{1}{2}v$ just by a permutation of coordinates. This permutation is always even by the following:

Claim. *Let A and B be finite sets of the same cardinality. Then the cardinality of the symmetric sum $A \vartriangle B$ is even.*

The orientations of the hyperplanes H_u and H_v are fixed by the condition that $\frac{1}{2}u + \frac{1}{2}v \in \operatorname{relint}\theta$. Thus, H_u and H_v are completely determined by the set of intervals that H_u cuts once. Let $A \subseteq \{1, \dots, j\}$ be the set of indices of intervals $I_1, \dots, I_h$ that H_u intersects once, and let $B \subseteq \{1, \dots, j\}$ be the same set for H_v. Then Ψ is a composition of a multiple of $A \vartriangle B$ transpositions and, hence, an even permutation. This means that all the local degrees (± 1's) in the sum (19) are of the same sign, and consequently $a = \pm\binom{j}{(j+1)/2}$.
Now, since d is even the equality (20) implies that

$$a \cdot \zeta = 4b \cdot \zeta.$$

Thus, if $\mathfrak{o}(g)$ is a coboundary a is divisible by 4. In the case $j = 2^t + 1$ where $t \geq 2$, and $d = 3 \cdot 2^{t-1} + 2$ the Kummer criterion implies that the binomial coefficient $\binom{j}{(j+1)/2}$ is divisible by 2 but *not* by 4. Hence, $\mathfrak{o}(g)$ is not a coboundary and a $\mathfrak{S}_2^{\pm}$-equivariant map $X_{d,2} \longrightarrow S(W_2 \oplus U_2^{\oplus j})$ whose restriction to $X_{d,2}^{>1}$ is $\mathfrak{S}_2^{\pm}$-homotopic to $\nu \circ \psi_{\mathcal{M}}|_{X_{d,2}^{>1}}$ does not exist.

This concludes the final instance (iii) of Theorem 1.5. □

4.4 Proof of Theorem 1.6

We prove both instances of the Ramos conjecture $\Delta(2,3) = 5$ and $\Delta(4,3) = 10$ using Theorem 1.4. Thus in order to prove that

- $\Delta(2,3) = 5$ it suffices to show that the number of non-equivalent 1-equiparting matrices of size $3 \times 2 \cdot 2^3$ is odd, Proposition 4.8;
- $\Delta(4,3) = 10$ it suffices to show that the number of non-equivalent 2-equiparting matrices of size $3 \times 4 \cdot 2^3$ is also odd, Enumeration 4.9.

Consequently we turn our attention to 3-bit Gray codes. It is not hard to see that the following lemma holds.

Lemma 4.7. *Let $c_1 \in \{0,1\}^3$ be a choice of first column.*

(i) *There are 18 different 3-bit Gray codes $A = (c_1, c_2, \ldots, c_8) \in \{0,1\}^{3\times 8}$ that start with c_1. They have transition counts $(3,2,2)$, $(3,3,1)$, or $(4,2,1)$.*

(ii) *There are 3 equivalence classes of Gray codes that start with with c_1. The three classes can be distinguished by their transition counts.*

Proof. (i): Starting at a given vertex of the 3-cube, there are precisely 18 Hamiltonian paths. This can be seen directly or by computer enumeration.

(ii): Follows directly from (i), as all equivalence classes have size 6: If $c_1 = (0,0,0)^t$ then all elements in a class are obtained by permutation of rows. For other choices of c_1, they are obtained by arbitrary permutations of rows followed by the "correct" row bit-inversions to obtain c_1 in the first column. □

Proposition 4.8. *There are 13 non-equivalent 1-equiparting matrices that are of size $3 \times (2 \cdot 2^3)$.*

Proof. Let $A = (A_1, A_2)$ be a 1-equiparting matrix. This means that both A_1 and A_2 are 3-bit Gray codes and the last column of A_1 is equal to the first column of A_2. In addition, the transition counts cannot exceed 5 and must sum up to 14. Having in mind that A is a 1-equiparting matrix it follows that A must have transition counts $\{5,5,4\}$. Hence two of its rows must have transition count 5 and one row must have transition count 4. In the following a *realization* of transition counts is a Gray code with the prescribed transition counts.

Since we are counting 1-equiparting matrices up to equivalence we may fix the first column of A, and hence first column of A_1, to be $(0,0,0)^t$ and choose for A_1 one of the matrices from each of the 3 classes of 3-bit Gray codes described in Lemma 4.7(ii).

If A_1 has transition counts $(3,2,2)$, i.e., the first row has transition count 3 while remaining rows have transition count 2, then its last column is $(1,0,0)^t$. The

next Gray code A_2 in the matrix a can have transition counts $(2,3,2)$, $(2,2,3)$, or $(1,3,3)$, each having 2 realizations A_2, each with first column $(1,0,0)^t$.
If A_1 has transition $(3,3,1)$, then its last column is $(1,1,0)^t$. The Gray code A_2 can have transition counts $(2,2,3)$, having 2 realizations, or $(1,2,4)$, having 1 realization, or $(2,1,4)$, having 1 realization, each with first column $(1,1,0)^t$.
If A_1 has transition counts $(4,2,1)$, then its last column is $(0,0,1)^t$. The Gray code A_2 can have transition counts $(1,2,4)$, having 1 realization, or $(1,3,3)$, having 2 realizations, each with first column $(0,0,1)^t$.
In total we have $6+4+3=13$ non-equivalent 1-equiparting matrices $A=(A_1,A_2)$. □

Enumeration 4.9. *There are* 2015 *non-equivalent* 2*-equiparting matrices that are of size* $3\times 4\cdot 2^3$.

Proof. Using Lemma 4.7 we enumerate non-equivalent 2-equiparting matrices by computer. Let $A=(A_1,A_2,A_3,A_4)$ be a 2-equiparting matrix. It must have transition counts $\{10,10,8\}$. Similarly as above, A is constructed by fixing the first column to be $(0,0,0)^t$ and A_1 to be one representative from each of the 3 classes of Gray codes. Then all possible Gray codes for A_2,A_3,A_4 are checked, making sure that the last column of A_i is equal to the first column of A_{i+1} and that the transition counts of $A_1,\dots,A_4$ sum up to $\{10,10,8\}$. This leads to 2015 possibilities. □

This concludes the proof of Theorem 1.6.

Remark 4.10. By means of a computer we were able to calculate the number $N(j,k,d)$ of non-equivalent ℓ-equiparting matrices for several values of $j\geq 1$ and $k\geq 3$, where $d=\lceil\frac{2^k-1}{k}j\rceil$ and $\ell=dk-(2^k-1)j$. See Table 1.

Number $N(j,k,d)$ of non-equiv ℓ-equiparting matrices given $j\geq 2$, and $k\geq 3$.				
j	k	ℓ	d	$N(j,k,d)$
2	3	1	5	13
3	3	0	7	60
4	3	2	10	2015
5	3	1	12	35040
6	3	0	14	185130
7	3	2	17	7572908
8	3	1	19	132909840
9	3	0	21	732952248
1	4	1	4	16
2	4	2	8	37964

Table 1: Number $N(j,k,d)$ of non-equivalent ℓ-equiparting matrices given $j\geq 2$ and $k\geq 3$, where $d=\lceil\frac{2^k-1}{k}j\rceil$ and $\ell=dk-(2^k-1)j$.

References

[1] David Avis, *Non-partitionable point sets*, Inform. Process. Letters 19 (1984), no. 3, 125–129.

[2] Anders Björner and Günter M. Ziegler, *Combinatorial stratification of complex arrangements*, J. Amer. Math. Soc. 5 (1992), 105–149.

[3] Pavle V. M. Blagojević, Florian Frick, Albert Haase, and Günter M. Ziegler, *Topology of the Grünbaum–Hadwiger–Ramos hyperplane mass partition problem*, Preprint, 27 pages, February 2015; arXiv:1502.02975.

[4] Pavle V. M. Blagojević, Wolfgang Lück, and Günter M. Ziegler, *Equivariant topology of configuration spaces*, J. Topology 8 (2015), 414–456.

[5] Pavle V. M. Blagojević and Günter M. Ziegler, *The ideal-valued index for a dihedral group action, and mass partition by two hyperplanes*, Topology Appl. 158 (2011), no. 12, 1326–1351.

[6] ———, *Convex equipartitions via equivariant obstruction theory*, Israel J. Math. 200 (2014), 49–77.

[7] Tammo tom Dieck, *Transformation Groups*, Studies in Mathematics, vol. 8, Walter de Gruyter, Berlin, 1987.

[8] Aleksandra S. Dimitrijević Blagojević, *A mass partition problem in* $\mathbf{R}^4$, Mat. Vesnik 61 (2009), 61–69.

[9] Ralph Fox and Lee Neuwirth, *The braid groups*, Math. Scandinavica 10 (1962), 119–126.

[10] Branko Grünbaum, *Partitions of mass-distributions and of convex bodies by hyperplanes*, Pacific J. Math. 10 (1960), no. 4, 1257–1261.

[11] Hugo Hadwiger, *Simultane Vierteilung zweier Körper*, Arch. Math. 17 (1966), no. 3, 274–278.

[12] Ernst E. Kummer, *Über die Ergänzungssätze zu den allgemeinen Reciprocitätsgesetzen*, J. Reine Angew. Math. 44 (1852), 93–146.

[13] Peter Mani-Levitska, Siniša Vrećica, and Rade T. Živaljević, *Topology and combinatorics of partitions of masses by hyperplanes*, Adv. Math. 207 (2006), no. 1, 266–296.

[14] Enrique Outerelo and Jesús M. Ruiz, *Mapping degree theory*, vol. 108, American Mathematical Society, 2009.

[15] Edgar A. Ramos, *Equipartition of mass distributions by hyperplanes*, Discrete Comput. Geom. 15 (1996), no. 2, 147–167.

[16] Siniša Vrećica and Rade T. Živaljević, *A proof of an equipartition result of Ramos*, Preprint, 8 pages, August 6, 2015; arXiv:1508.01552v1.

[17] Rade T. Živaljević, *Computational topology of equipartitions by hyperplanes*, Top. Meth. Nonlinear Analysis 45 (2015), 63–90.

Pavle V. M. Blagojević
Math. Institut SANU
Knez Mihailova 36
11001 Beograd
Serbia
and
Institute of Mathematics
Freie Universität Berlin
Arnimallee 2
14195 Berlin
Germany
pavleb@mi.sanu.ac.rs

Florian Frick
Department of Mathematics
Cornell University
Ithaca
NY 14853
USA
ff238@cornell.edu

Albert Haase
Institute of Mathematics
Freie Universität Berlin
Arnimallee 2
14195 Berlin
Germany
a.haase@fu-berlin.de

Günter M. Ziegler
Institute of Mathematics
Freie Universität Berlin
FU Berlin
Arnimallee 2
14195 Berlin
Germany
ziegler@math.fu-berlin.de

Quadratic and Symmetric Bilinear Forms on Modules with Unique Base over a Semiring

Zur Izhakian, Manfred Knebusch, Louis Rowen

Received: September 4, 2015
Revised: December 3, 2015

Communicated by Ulf Rehmann

Abstract. We study quadratic forms on free modules with unique base, the situation that arises in tropical algebra, and prove the analog of Witt's Cancelation Theorem. Also, the tensor product of an indecomposable bilinear module (U, γ) with an indecomposable quadratic module (V, q) is indecomposable, with the exception of one case, where two indecomposable components arise.

2010 Mathematics Subject Classification: Primary 15A03, 15A09, 15A15, 16Y60; Secondary 14T05, 15A33, 20M18, 51M20
Keywords and Phrases: Semirings, (semi)modules, bilinear forms, quadratic forms, symmetric forms, orthogonal decomposition.

1. Overview

This paper is part of a program to understand the theory of quadratic forms over the max-plus algebra and related semirings that arise in several mathematical contexts. Our motivation comes from two sources, tropical mathematics and real algebra, which interact with each other. Since the first area is still in its nascent stage, for the reader's convenience, we provide a short overview of this mathematics and related subjects.

Consider the field $\mathbb{K}$ of *Puiseux series* over an algebraically closed field F of characteristic 0. The elements of $\mathbb{K}$ are of the form

$$f = \sum_{\tau \in \mathbb{Q}} c_\tau t^\tau,$$

where $c_\tau \in F$ and the powers of t are taken over well-ordered subsets of $\mathbb{Q}$. (In the literature one often takes $\mathbb{R}$ instead of $\mathbb{Q}$.)

Define the *order valuation* $v : \mathbb{K} \to \mathbb{Q}$ by

$$v(f) \ := \min\{\tau \in \mathbb{Q}_{\geq 0} \ : \ c_\tau \neq 0\}$$

for which the dominant term in f becomes $c_{v(f)}t^{v(f)}$ as $t \to 0$. Then v is a valuation, with residue field F, with respect to which $\mathbb{K}$ is complete and thus Henselian. By Hensel's lemma, $\mathbb{K}$ also is algebraically closed, and thus elementarily equivalent to F. Applying v takes us from $\mathbb{K}$ to the ordered group $\mathbb{Q}$, which can be viewed as a "max-plus" semiring (taking $-v$ instead of v), whose operations are "+" for multiplication and "sup" for addition. This process, called *tropicalization*, is explained in [15, 29]. The point of tropicalization is to simplify the combinatorics in algebraic geometry and linear algebra, and there has been considerable success in this direction in enumerative geometry.

One can tropicalize structures arising in linear algebra, such as quadratic forms, simply by replacing the classical addition and multiplication by the max-plus operations respectively, but then the classical theory does not go through because our new addition (max) does not have negatives.

Other important (non-tropical) semirings, where our below theory is relevant, occur in real algebra, such as the positive cone of an ordered field [4, p. 18] or a partially ordered commutative ring [5, p. 32]. A further application can be found in the algebra of groups over a splitting field, as described briefly at the end of this overview.

Recall that a (commutative) *semiring* is a set R equipped with addition and multiplication, such that both $(R, +, 0)$ and $(R, \cdot, 1)$ are abelian monoids with elements $0 = 0_R$ and $1 = 1_R$ respectively, and multiplication distributes over addition in the usual way. In other words, R satisfies all the properties of a commutative ring except the existence of negation under addition. We call a semiring R a *semifield*, if every nonzero element of R is invertible; hence $R \setminus \{0\}$ is an abelian group.

As in the classical theory, one considers bilinear and quadratic forms defined on (semi)modules over a semiring R, often a "supersemifield," in order to obtain more sophisticated "trigonometric" information, cf. [24, §2, §3].

On one hand, these semirings lack negation, thereby playing havoc even with the notion of the underlying bilinear form of a quadratic form. On the other hand, they have the pleasant property that free modules have "unique base," cf. Definition 1.2. Thus, our overall object is to classify quadratic forms over free modules having unique base, with applications to the supertropical setting.

For the reader's convenience, we recall some terminology and results from [22, §1-§4]. A *module* V over R (sometimes called a *semimodule*) is an abelian monoid $(V, +, 0_V)$ equipped with a scalar multiplication $R \times V \to V$, $(a, v) \mapsto av$, such that exactly the same axioms hold as customary for modules over a ring: $a_1(bv) = (a_1b)v$, $a_1(v + w) = a_1v + a_1w$, $(a_1 + a_2)v = a_1v + a_2v$, $1_R \cdot v = v$, and $0_R \cdot v = 0_V = a_1 \cdot 0_V$ for all $a_1, a_2, b \in R$, $v, w \in V$. We write 0 for both 0_V and 0_R, and 1 for 1_R.

When considering modules over semifields, one encounters several versions of "base," as studied in depth in [21, §4 and §5.3]. Here we take the standard categorical version, and call an R-module V *free*, if there exists a family $(\varepsilon_i \mid i \in I)$ in V such that every $x \in V$ has a unique presentation $x = \sum_{i \in I} x_i\varepsilon_i$ with scalars

$x_i \in R$ and only finitely many x_i nonzero, and we call $(\varepsilon_i \mid i \in I)$ a *base* of the R-module V. Any free module with a base of n elements is clearly isomorphic to R^n, under the map $\sum_{i=1}^{n} x_i\varepsilon_i \mapsto (x_1, \dots, x_n)$.

Bilinear forms on V are defined in the obvious way, [21].

DEFINITION 1.1. *For any module V over a semiring R, a* quadratic form *on V is a function $q : V \to R$ with*

$$q(ax) = a^2 q(x) \tag{1.1}$$

for any $a \in R$, $x \in V$, together with a symmetric bilinear form $b : V \times V \to R$ (not necessarily uniquely determined by q) such that for any $x, y \in V$

$$q(x + y) = q(x) + q(y) + b(x, y). \tag{1.2}$$

Every such bilinear form b will be called a companion *of q, and the pair (q, b) will be called a* quadratic pair *on V. We also call V a* quadratic module.

In this generality, it is difficult to describe quadratic forms adequately on free modules over an arbitrary semiring. However, our task becomes more manageable when we introduce the following condition.

DEFINITION 1.2. *An R-*module with unique base *is a free R-module V in which any two bases $\mathfrak{B}$, $\mathfrak{B}'$ are projectively the same, i.e., we obtain the elements of $\mathfrak{B}'$ from those of $\mathfrak{B}$ by multiplying by units of R.*

Although this never happens for free modules of rank ≥ 2 over a ring, it turns out to be quite common in the context of tropical algebra (and also often in real algebra, as noted in Example 2.4.d).

Our main result, in §5, is an analog of Witt's cancelation theorem:

THEOREM 5.9.*If W_1, W_1', W_2, W_2' are finitely generated quadratic or bilinear modules with unique base such that $W_1 \cong W_1'$ and $W_1 \perp W_2 \cong W_2 \perp W_2'$, then $W_2 \cong W_2'$ (where $\cong$ means "isometric").*

It actually is given in more general terms, where W_2 needs not be finitely generated.

When R is a ring, then a quadratic form q has just one companion, namely,

$$b(x, y) := q(x + y) - q(x) - q(y),$$

but if R is a semiring that cannot be embedded into a ring, this usually is not the case, and it is a major concern of quadratic form theory over semirings to determine all companions of a given quadratic form $q : V \to R$.

The first step in classifying quadratic forms is [22, Propositions 4.1 and 4.2], which lets us write a quadratic form q as the sum $q = \kappa + \rho$, where κ is *quasilinear* (and unique) in the sense that $\kappa(x + y) = \kappa(x) + \kappa(y)$, and ρ is *rigid* in the sense that it has a unique companion. Quasilinearity of a quadratic form q implies that, for any vector $x = \sum_{i\in I} x_i\varepsilon_i$ in V,

$$q(x) = \sum_{i\in I} x_i^2 q(\varepsilon_i), \tag{1.3}$$

i.e., q has *diagonal form* with respect to the base $(\varepsilon_i : i \in I)$.
Quasilinear forms follow aspects of the classical theory of quadratic forms, and satisfy a Cauchy-Schwartz inequality given in [24]. On the other hand, by [22, Theorem 3.5], the rigid forms are precisely those with $q(\varepsilon_i) = 0$ for all $i \in I$.
Our ultimate object being to classify quadratic forms over free modules with unique base, in this paper we study quadratic forms in terms of orthogonal decompositions of such forms into indecomposable forms, and then build them up again via tensor products of two symmetric bilinear forms and of a symmetric bilinear form with a quadratic form.
Let us turn now to the tools needed in proving Theorem 5.9.

1.1. Partial quasilinearity. We seldom require quasilinearity in its entirety, but the following partial version plays a major role in our consideration of orthogonal decompositions of quadratic modules.

Definition 1.3. *Given subsets S and T of V, we say that q is* quasilinear on $S \times T$ *if*

$$q(x + y) = q(x) + q(y).$$

for all $x \in S$, $y \in T$.

The following helpful fact is a special case of [22, Lemma 1.18]. (We write $S + S'$ for $\{s + s' : s \in S, s' \in S'\}$.)

Lemma 1.4. *Let S, S', T be subsets of V. If q is quasilinear on $S \times T$, $S' \times T$ and $S \times S'$, then q is quasilinear on $(S + S') \times T$.*

1.2. Disjoint orthogonality. In §3 we develop the notion of *(disjoint) orthogonality* of two given submodules W_1 and W_2 of a *quadratic R-module* (V, q) (endowed with a fixed quadratic form q), which means that $W_1 \cap W_2 = \{0\}$ and q is partially quasilinear on $W_1 \times W_2$. (Note that there is no direct reference to an underlying symmetric bilinear form.) When V has unique base, we look for orthogonal decompositions $V = W_1 \perp W_2$, and more generally $V = \underset{i \in I}{\perp} W_i$, where the W_i are *basic submodules* of V, i.e., are generated by subsets of a base $\mathfrak{B}$ of V.
We can choose a companion b of q (called "quasiminimal" companion) adapted to the notion of disjoint orthogonality, and then have an equivalence relation on the set $\mathfrak{B}$ at hands, which is generated by the pairs $(\varepsilon, \varepsilon')$ in $\mathfrak{B}$ with $\varepsilon \neq \varepsilon'$, $b(\varepsilon, \varepsilon') \neq 0$. By the use of this equivalence relation the indecomposable basic submodules of V (in the sense of disjoint orthogonality) can be described as follows.

Theorem 3.8.*Let $\{\mathfrak{B}_k \mid k \in K\}$ denote the set of equivalence classes in $\mathfrak{B}$ and, for every $k \in K$, let W_k denote the submodule of V having base $\mathfrak{B}_k$.*

(a) *Then every W_k is an indecomposable basic submodule of V and*

$$V = \underset{k \in K}{\perp} W_k.$$

(b) *Every indecomposable basic submodule U of V is contained in W_k, for some $k \in K$ uniquely determined by U.*

(c) *The modules W_k, $k \in K$, are precisely all the indecomposable basic orthogonal summands of V.*

In §4 we develop the analogous notion of disjoint orthogonality in a *bilinear R-module* (V, b) with respect to a fixed symmetric bilinear form b on V, and we show:

THEOREM 4.9. *If b is a quasiminimal companion of a a quadratic module (V, q), then the indecomposable components of (V, q) coincide with the indecomposable components of (V, b).*

In §5, these decomposition theories yield the desired analog (Theorem 5.9) of Witt's cancelation theorem.

1.3. TENSOR PRODUCTS. The last two sections of the paper are devoted to tensor products. Whereas tensor products of modules over general semirings can be carried out in analogy with the usual classical construction over rings, it requires the use of congruences, resulting in some technical issues dealt with in [7, Chap. 16], for example. But for free modules with unique base the construction can be carried out easily, since then one does not need to worry about well-definedness.

In §6 we construct the tensor product of two free bilinear R-modules over any semiring R, in analogy to the case where R is a ring, cf. [8, §2], [26, I, §5]. We then take the tensor product of a free bilinear R-module $U = (U, \gamma)$ with a free quadratic R-module $V = (V, q)$. A new phenomenon occurs here, in contrast to the theory over rings. It is necessary first to choose a so-called *balanced companion* b of q, which always exists, cf. [22, §1], but which usually is not unique. We then define the tensor product $U \otimes_b V$, depending on b, by choosing a so-called *expansion* $B : V \times V \to R$ of the quadratic pair (q, b) which is a (not necessarily symmetric) bilinear form B with

$$B(x, x) = q(x), \qquad B(x, y) + B(y, x) = b(x, y)$$

for all $x, y \in V$, cf. [22, §1] and then proceed essentially as in the case of rings, e.g. [26, Definition 1.51], [8, p. 51][1]. The resulting quadratic form $\gamma \otimes_b q$ does not depend on the choice of B but often depends on the choice of b. This is apparent already in the case $\gamma = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, where the matrix b is stored in the quadratic polynomial $\gamma \otimes_b q$, cf. Example 6.8 below.

In §7 we turn to the indecomposability of tensor products. For convenience, we assume that $R \setminus \{0\}$ is closed under multiplication and addition, implying by Theorem 2.3 that all free R-modules have unique base.

After obtaining partial results along the way, we arrive at the main result of this section, Theorem 7.16, which states that, discarding trivial situations and

[1]When R is a ring the "b" in the tensor product is not specified since q has only one companion.

excluding some pathological semirings, the tensor product of an indecomposable bilinear module (U, γ) with an indecomposable quadratic module (V, q) is again indecomposable, with the exception of one case, where two indecomposable components arise.

1.4. Applications. The remainder of this introduction discusses how quadratic forms over modules with unique base over semirings arise naturally in various contexts in mathematics. (The reader could skip directly on to the main theoretical results of this paper.)

1.4.1. *Quadratic forms over rings.* Supertropical semirings, to be defined below (cf. [25, 22]), establish a class of semirings over which every free module has a unique base. There is a way to pass from a quadratic form on a free module over a (commutative) ring R to quadratic forms on free modules over a supertropical semiring U. To explain this, we sketch the notion of *supertropicalization* of a quadratic form $q : V \to R$, obtained by a so-called *supervaluation* $\varphi : R \to U$.
An *m-valuation* (= monoid valuation) on a ring R is a map $v : R \to M$ from R to a totally ordered abelian monoid $M = (M, \cdot\ , \leq)$, containing an absorbing element $0 = 0_M$ $(0 \cdot x = x \cdot 0 = 0)$ with $0 \leq x$ for all $x \in M$, which satisfies the following rules:

$$v(0) = 0, \quad v(1) = 1, \quad v(xy) = v(x)v(y),$$

and

$$v(x + y) \leq \max\{v(x), v(y)\} \tag{1.4}$$

for all $x, y \in M$. When $\Gamma := M \setminus \{0\}$ is a group, we call the m-valuation $v : R \to M$ a *valuation*. These are exactly the valuations as defined by Bourbaki [3] and studied, e.g., in [14] and [27, Ch. I], except that for Γ we have chosen the multiplicative notation instead of the additive notation. In this case $v^{-1}(0)$ is a prime ideal of R [loc. cit.]. When R is a field this forces $v^{-1}(0) = \{0\}$, and we return to Krull valuations.
Given an m-valuation $v : R \to M$, we equip M with the additive operation defined as

$$a + b := \max\{a, b\},$$

which makes M a *bipotent semiring*, i.e., a semiring M' in which $a + b \in \{a, b\}$ for all $a, b \in M'$. Conversely any bipotent semiring M' has a natural total order given by

$$a < b \quad \Leftrightarrow \quad a + b = b,$$

and can be viewed as a totally ordered abelian monoid with an absorbing element $0_{M'}$. Therefore, totally ordered monoids M with zero can be referred to as bipotent semirings (or bipotent semifields when $M \setminus \{0\}$ is a group). Viewed in this way, rule (1.4) reads

$$v(x + y) \leq v(x) + v(y). \tag{1.5}$$

This brings us into the realm of semirings. A semiring U is called *supertropical* if the following conditions hold:

- $e := 1_U + 1_U$ is idempotent (i.e., $2 \times 1 = 4 \times 1$),
- the *ghost ideal* $M = eU$ is a bipotent semiring,
- addition is defined in terms of the *ghost map* $a \mapsto ea$ and the ordering of M, as follows:

$$a + b = \begin{cases} a & \text{if } ea < eb; \\ b & \text{if } eb < ea; \\ ea & \text{if } ea = eb. \end{cases} \tag{1.6}$$

In particular $ea = 0$ implies $a = 0$ (take $b = 0$ in (1.6)). The elements of eU are called *ghost elements* and those of $U \setminus eU$ are called *tangible elements.* The zero element is regarded both as tangible and ghost. See [17, 18, 25] for the ideas behind this terminology.

A *supervaluation* on a ring R is a multiplicative map $\varphi : R \to U$ sending R into a supertropical semiring, such that $\varphi(0) = 0$, $\varphi(1) = 1$, and

$$e\varphi(x + y) \leq e\varphi(x) + e\varphi(y)$$

for all $x, y \in R$. The map $v := e\varphi : R \to M$, $x \mapsto e\varphi(x)$, is then an m-valuation, which as we say is *covered* by φ. For any given m-valuation $v : R \to M$, there usually is an extended hierarchy of supervaluations $\varphi : R \to U$ covering v (with $U \supset M$, $eU = M$, U varying) studied in [17, 18].

The supertropicalizations of a quadratic form $q : V \to R$ on a free R-module V are constructed by using a supervaluation $\varphi : R \to U$ as follows. We choose an ordered base $\mathcal{L}$ of V, say $\mathcal{L} = \{v_i : i \in I\}$ with $I = \{1, \dots, n\}$, and write q as a homogenous polynomial of degree 2

$$q\biggl(\sum_{i=1}^{n} x_i v_i\biggr) = \sum_{i=1}^{n} \alpha_i x_i^2 + \sum_{i<j} \beta_{ij} x_i x_j, \tag{1.7}$$

with $\alpha_i = q(v_i)$, $\beta_{ij} = b(v_i, v_j)$, where b is the (unique) companion of q. We denote by U^n the free U-module consisting of all n-tuples in U. Let $\{\varepsilon_1, \dots, \varepsilon_n\}$ be the standard base of U, where each ε_i has i-th coordinate 1 and all other coordinates 0. Using a new set of variables $\lambda_1, \dots, \lambda_n$, we define

$$q^{\varphi}\biggl(\sum_{i=1}^{n} \lambda_i \varepsilon_i\biggr) := \sum_{i=1}^{n} \varphi(\alpha_i) \lambda_i^2 + \sum_{i<j} \varphi(\beta_{ij}) \lambda_i \lambda_j \tag{1.8}$$

by applying φ to the coefficients of the polynomial (1.7).

We write $(U^{(I)}, q^{\varphi})$, or $(\widetilde{V}, \tilde{q})$ for short, for the *supertropicalization* of the quadratic module (V, q) with respect to the base $\mathcal{L}$. Since every U-module has a unique base, cf. §2, the base $\{\varepsilon_i : i \in I\}$ of $\widetilde{V}$ is unique up to permuting the ε_i and multiplying them by units of U (which are the invertible tangible elements of U). That is, the base $\mathcal{L}$ of V becomes "frozen" in the free quadratic module $(\widetilde{V}, \tilde{q})$ obtained from (V, q) by a kind of "degenerate scalar extension" $\varphi : R \to U$. {φ is multiplicative, but respects addition only weakly.} This central fact motivates our interest in supertropicalization.

One reason that we work with m-valuations in general, instead of just valuations covered by supervaluations, is that m-valuations which are not valuations often arise naturally in the context of commutative algebra as described in the paper [11] of Harrison and Vitulli. They construct so-called "V-valuations" (there named "formally finite" V-valuations). This construction has been complemented later by D. Zhang with somewhat dual "V^0-valuations" [33]. These constructions have been revised in [19, §1-§3], showing that any m-valuation on a ring can be coarsened both to a V-valuation and to a V^0-valuation, and also to a valuation in a minimal way.

In [12] Harrison and Vitulli, pursuing their idea of "infinite primes" (in the sense of classical number theory) from [11], construct $\mathbb{C}$-valued places on a field by a somewhat similar method. This construction has been extended by Valente and Vitulli in [31] to "preplaces" on a ring R, which are interpreted in [19] as multiplicative maps $\chi : R \to R'$ to a bipotent semiring R' such that $\chi(0) = 0$, $\chi(1) = 1$, and

$$\chi(x + y) \le c(\chi(x) + \chi(y))$$

for all $x, y \in R$, where c is a unit of R'. Such a map χ provides various supervaluations $\varphi : R \to U$ that cover V-valuations $v : R \to eU$ [19, §4]. Since the multiplicative monoids $eU \setminus \{0\}$ are cancellative, these V-valuations are true valuations. By a related method, supervaluations arise that cover V^0-valuations, which again are true valuations.

Although not all supervaluations can be constructed in this way, at least we gain a rich stock of m-valuations and supervaluations on a ring. Facing a problem on quadratic forms over a ring R, it may be a piece of art to address an appropriate supervaluation which fits best the supertropical framework. Much space is left for further study in this research direction.

1.4.2. *A surprise.* In an earlier version of this paper we considered quadratic forms over supertropical semirings, knowing already from [22, Theorem 0.9] that a free module over these semirings has unique base, and we obtained the results in §3-§7 for such quadratic forms. Only later did we realize that these results go through for any semiring R over which all free modules have unique base. As a consequence, supertropical semirings hardly appear explicitly in §3-§7. This paves the way for an extra application, which we now describe. Namely, take an algebra A with a bilinear form, whose orthogonal base generates a natural proper semiring of A.

1.4.3. *Table algebras.* A classical example is the set of characters of a finite group G over a field whose characteristic does not divide $|G|$; since the sum (resp. product) of characters is the character of the direct sum (resp. tensor product) of their underlying representations, we can restrict to the semiring of characters, which is a free module over $\mathbb{N}_0$. A similar situation arises for the center of the group algebra, which is a free module whose base is comprised of the sums of elements from conjugacy classes. These algebras have been generalized by Hoheisel [13] and Arad-Blau [1] as explained in the fine survey by Blau [2], where he defines Hoheisel algebras and table algebras. These have a distinguished base $\mathcal{L}$ that spans the sub-semialgebra A^+ that they generate over $\mathbb{R}^+$, so again A^+ is a free module over $\mathbb{R}^+$ (with unique base $\mathcal{L}$), and a natural framework in which to build quadratic forms.

2. R-modules with unique base and their basic submodules

We assume throughout this paper that V is a free R-module with unique base $\mathfrak{B}$. Accordingly, we begin by examining this property.

Remark 2.1. *Any change of base of the free module R^n is attained by multiplication by an invertible $n \times n$ matrix, so having unique base is equivalent to every invertible matrix in $M_n(R)$ being a generalized permutation matrix.*

Our interest in these modules stems from the following key fact.

Theorem 2.2 ([21, Corollary 5.25] and [22, Theorem 0.9]). *If R is a supertropical semiring, then every free R-module has unique base.*

More generally, one may ask, "What conditions on the semiring R guarantee that R^n has unique base, or equivalently, that every invertible matrix is a generalized permutation matrix?" The matrix question was answered in [30] and [6]. In their terminology, an "antiring" is a semiring R such that $R \setminus \{0\}$ is closed under addition. We prefer the terminology "lacks zero sums," since this property holds also for sums of squares in a real closed field, and "antiring" does not seem appropriate in that context.
Tan and Dolžna-Oblak classify the invertible matrices over these rings lacking zero sums. These are just the generalized permutation matrices when $R \setminus \{0\}$ also is closed under multiplication, which they call "entire" (the case in tropical mathematics), and more generally by [6, Theorem 1] (as interpreted in Theorem 2.5) when R is indecomposable, i.e., not isomorphic to a direct product $R_1 \times R_2$ of semirings.

Theorem 2.3 (cf. [6, §2, Corollary 3], an alternative proof given below). *If the set $R \setminus \{0\}$ is closed under addition and multiplication (i.e., $a + b = 0 \Rightarrow a = b = 0$, $a \cdot b = 0 \Rightarrow a = 0$ or $b = 0$), then every free R-module has unique base.*

In view of Remark 2.1, Theorem 2.3 follows from Dolžan and Oblak [6, §2, Corollary 3] using matrix arguments within a wider context extending work of Tan [30, Proposition 3.2], which in turn relies on Golan's book on semirings [9, Lemma 19.4].

Example 2.4. *Here are some instances where $R \setminus \{0\}$ is closed under addition and multiplication.*

a) *The "Boolean semifield" $\mathbb{B} = \{-\infty, 0\}$ (and thus subalgebras of algebras that are free modules over $\mathbb{B}$). This shows that our results pertain to "$\mathcal{F}_1$-geometry."*
b) *Rewriting the Boolean semifield instead as $\mathbb{B} = \{0, 1\}$ where $1 + 1 = 1$, one can generalize it to $\{0, 1, \dots, q\}$ $L = [1, q] := \{1, 2, \dots, q\}$ the "truncated semiring without 0" of* [23, Example 2.14], *where "$a + b$" is defined to be the minimum of their sum and q.*

c) *Function semirings, polynomial semirings, and Laurent polynomial semirings over these semirings.*
d) *If F is a formally real field, i.e. -1 is not a sum of squares in F, then the subsemiring $R = \Sigma F^2$, consisting of all sums of squares in F, lacks zero sums. In fact R is a semifield; the inverse of a sum of squares*

$$a = x_1^2 + \cdots + x_r^2 \quad \textit{is} \quad a^{-1} = \left(\frac{x_1}{a}\right) + \cdots + \left(\frac{x_r}{a}\right)^2.$$

Other than the trivial fact that every free R-module of rank 1 has unique base, all examples known to us of modules with unique base stem from Theorem 2.5, which is essentially [6, Theorem 1]:

Theorem 2.5 ([6, Theorem 1]). *Assume that R is an indecomposable semiring lacking zero sums. Then every free R-module has unique base.*

We now reprove Theorem 2.3 by a simple matrix-free argument in preparation for a reproof of the more general Theorem 2.5.

Proof of Theorem 2.3. Let V be a free R-module and $\mathfrak{B}$ a base of V. If $x \in V \setminus \{0\}$ is given, we have a presentation

$$x = \sum_{i=1}^{r} \lambda_i x_i$$

with $x_i \in \mathfrak{B}$ and $\lambda_i \in R \setminus \{0\}$. We call the set $\{x_1, \dots, x_r\} \subset \mathfrak{B}$ the *support* of x with respect to $\mathfrak{B}$ and denote this set by $\operatorname{supp}_{\mathfrak{B}}(x)$. Note that if $x, y \in V \setminus \{0\}$, then $x + y \neq 0$ and

$$\operatorname{supp}_{\mathfrak{B}}(x+y) = \operatorname{supp}_{\mathfrak{B}}(x) \cup \operatorname{supp}_{\mathfrak{B}}(y) \tag{2.1}$$

due to the assumption that $\lambda + \mu \neq 0$ for any $\lambda, \mu \in R \setminus \{0\}$. Also

$$\operatorname{supp}_{\mathfrak{B}}(\lambda x) = \operatorname{supp}_{\mathfrak{B}}(x) \tag{2.2}$$

for $x \in V \setminus \{0\}$, $\lambda \in R \setminus \{0\}$, due to the assumption that for $\lambda, \mu \in R \setminus \{0\}$ we have $\lambda\mu \neq 0$.
Now assume that $\mathfrak{B}'$ is a second base of V. Given $x \in \mathfrak{B}$, we have a presentation

$$x = \lambda_1 y_1 + \cdots + \lambda_r y_r$$

with $\lambda_i \in R \setminus \{0\}$ and distinct $y_i \in \mathfrak{B}'$. It follows from (2.1) and (2.2) that

$$\{x\} = \operatorname{supp}_{\mathfrak{B}}(x) = \operatorname{supp}_{\mathfrak{B}}(y_1) \cup \cdots \cup \operatorname{supp}_{\mathfrak{B}}(y_r).$$

This forces

$$\{x\} = \operatorname{supp}_{\mathfrak{B}}(y_1) = \cdots = \operatorname{supp}_{\mathfrak{B}}(y_r). \tag{2.3}$$

¿From this, we infer that $r = 1$. Indeed, suppose that $r \geq 2$. Then $y_1 = \mu_1 x$, $y_2 = \mu_2 x$ with $\mu_1, \mu_2 \in R \setminus \{0\}$. But this implies $\mu_2 y_1 = \mu_1 y_2$, a contradiction since y_1, y_2 are different elements of a base of V.
Thus $\{x\} = \operatorname{supp}_{\mathfrak{B}}(y)$ for a unique $y \in \mathfrak{B}'$, which means $y = \lambda x$ with $\lambda \in R \setminus \{0\}$. By symmetry we have a unique $z \in \mathfrak{B}$ and $\mu \in R \setminus \{0\}$ with $x = \mu z$. Then $x = \lambda\mu z$, whence $x = z$ and $\lambda\mu = 1$. Thus $\lambda, \mu \in R^*$ and $x \in R^* y$,

$y \in R^* x$. Of course, y runs through all of $\mathfrak{B}'$ if x runs through $\mathfrak{B}$, since both $\mathfrak{B}$ and $\mathfrak{B}'$ span the module V. $\square$

Proof of Theorem 2.5. Assume that $\mathfrak{B}$ and $\mathfrak{B}'$ are bases of V. Given $x \in \mathfrak{B}$, we write again

$$x = \lambda_1 y_1 + \cdots + \lambda_r y_r \tag{2.4}$$

with different $y_i \in \mathfrak{B}'$, $\lambda_i \in R \setminus \{0\}$. But now, instead of (2.3) we can only conclude that

$$\{x\} = \operatorname{supp}_{\mathfrak{B}}(\lambda_1 y_1) = \cdots = \operatorname{supp}_{\mathfrak{B}}(\lambda_i y_i). \tag{2.5}$$

Thus we have scalars $\mu_i \in R \setminus \{0\}$ such that

$$\lambda_i y_i = \mu_i x \qquad \text{for} \quad 1 \leq i \leq r. \tag{2.6}$$

Suppose that $r \geq 2$. Then we have for all $i, j \in \{1, \dots, r\}$ with $i \neq j$.

$$\mu_j \lambda_i y_i = \mu_j \mu_i x = \mu_i \mu_j x = \mu_i \lambda_j y_j.$$

Since the y_i are elements of a base, this implies $\mu_i \lambda_j = \mu_j \lambda_i = 0$ for $i \neq j$ and then

$$\mu_i \mu_j = 0 \qquad \text{for} \quad i \neq j. \tag{2.7}$$

On the other hand, we obtain from (2.4) and (2.6) that

$$x = \mu_1 x + \mu_2 x + \cdots + \mu_r x,$$

and then

$$1 = \mu_1 + \mu_2 + \cdots + \mu_r. \tag{2.8}$$

Multiplying (2.8) by μ_i and using (2.7), we obtain

$$\mu_i^2 = \mu_i. \tag{2.9}$$

Thus

$$R \cong R\mu_1 \times \cdots \times R\mu_r.$$

This contradicts our assumption that R is indecomposable.
We have proved that $r = 1$. Thus for every $x \in \mathfrak{B}$ there exist unique $y \in \mathfrak{B}'$ and $\lambda \in R$ with $x = \lambda y$. By the same argument as in the end of proof of Theorem 2.3, we conclude that $\mathfrak{B}$ is projectively unique. $\square$

Of course, if $R \setminus \{0\}$ is closed under multiplication, i.e., R has no zero divisors, then R is indecomposable. This also holds when R is supertropical (cf. [25, §3], [22, Definition 0.3]), since then for any two elements μ_1, μ_2 of R with $\mu_1 + \mu_2 = 1$ either $\mu_1 = 1$ or $\mu_2 = 1$. Thus, Theorem 2.5 generalizes both Theorems 2.2 and 2.3.
The following example reveals that Theorem 2.5 is the best we can hope for, in order to guarantee that every free R-module has unique base, as long as we stick to the natural assumption that R is a semiring lacking zero sums.

Example 2.6. *If R_0 is a semiring lacking zero sums, then $R := R_0 \times R_0$ also lacks zero sums. Put $\mu_1 = (1,0)$, $\mu_2 = (0,1)$. These are idempotents in R with $\mu_1\mu_2 = 0$ and $\mu_1 + \mu_2 = 1$. Now let V be a free R-module with base $\mathfrak{B} = \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n\}$, $n \geq 2$, choose a permutation $\pi \in S_n$, $\pi \neq 1$, and define*

$$\varepsilon_i' := \mu_1\varepsilon_i + \mu_2\varepsilon_{\pi(i)} \qquad (1 \leq i \leq n).$$

We claim that $\mathfrak{B}' := \{\varepsilon_1', \dots, \varepsilon_n'\}$ is another base of V.
Indeed, V is a free R_0-module with base $(\mu_i\varepsilon_j \mid 1 \leq i \leq 2,\ 1 \leq j \leq n)$. We have

$$\mu_1\varepsilon_i' = \mu_1\varepsilon_i, \qquad \mu_2\varepsilon_i' = \mu_2\varepsilon_{\pi(i)},$$

and thus $(\mu_i\varepsilon_j' \mid 1 \leq i \leq 2,\ 1 \leq j \leq n)$ is a permutation of this base over R_0, i.e., regarded as a set, the same base. Thus certainly $\mathfrak{B}'$ spans V as R-module. Given $x \in V$, let $x = \sum_1^n a_i\varepsilon_i'$ with $a_i \in R$. We have

$$a_i = a_{i1}\mu_1 + a_{i2}\mu_2 \qquad \text{with} \qquad a_{i1} \in R_0,\ a_{i2} \in R_0,$$

whence

$$x = \sum_{i=1}^n a_{i1}(\mu_1\varepsilon_i) + \sum_{i=1}^n a_{i2}(\mu_2\varepsilon_{\pi(i)}).$$

This shows that the coefficients $a_{i1}, a_{i2} \in R_0$ are uniquely determined by x, whence the coefficients $a_i \in R$ are also uniquely determined by x. Our claim is proved.
Since $\operatorname{supp}_{\mathfrak{B}}(\varepsilon_i')$ *has two elements if $\pi(i) \neq i$, $\mathfrak{B}'$ differs projectively from $\mathfrak{B}$. The base $\mathfrak{B}$ of the R-module V is not unique.*

3. Orthogonal decompositions of quadratic modules with unique base

Assume that V is an R-module equipped with a fixed quadratic form $q : V \to R$. We then call $V = (V, q)$ a *quadratic R-module.*

Definition 3.1.

(a) *Given two submodules W_1, W_2 of the R-module V, we say that W_1 is* disjointly orthogonal *to W_2, if $W_1 \cap W_2 = \{0\}$ and $q(x+y) = q(x)+q(y)$ for all $x \in W_1$, $y \in W_2$, i.e., q is quasilinear on $W_1 \times W_2$. (We say "orthogonal" for short, when it is clear a priori that $W_1 \cap W_2 = \{0\}$.)*

(b) *We write $V = W_1 \perp W_2$ if $V = W_1 \oplus W_2$ (as R-module) with W_1 disjointly orthogonal to W_2. We then call W_1 an* orthogonal summand *of W, and W_2 an* orthogonal complement *of W_1 in V.*

Caution. If $V = W_1 \perp W_2$, we may choose a companion b of q such that $b(W_1, W_2) = 0$, but note that it could well happen that the set of all $x \in V$ with $b(x, W_1) = 0$ is bigger than W_2, even if R is a semifield and $q|W_1$ is anisotropic (e.g., if q itself is quasilinear). Our notion of orthogonality does not refer to any bilinear form.

We now also define infinite orthogonal sums. This seems to be natural, even if we are originally interested only in finite orthogonal sums. Indeed, even if R is a semifield, a free R-module with finite base often has many submodules which are not finitely generated.

Definition 3.2. *Let $(V_i \mid i \in I)$ be a family of submodules of the quadratic module V. We say that V is* the orthogonal sum of the family (V_i), *and then write*

$$V = \underset{i\in I}{\perp} V_i,$$

if for any two different indices i, j the submodule V_i is disjointly orthogonal to V_j, and moreover $V = \bigoplus_{i\in I} V_i$.

N.B. Of course, then for any subset $J \subset I$, the module $V_J = \sum_{i\in J} V_i$ is the orthogonal sum of the subfamily $(V_i \mid i \in J)$; in short,

$$V_J = \underset{i\in J}{\perp} V_i.$$

We state a fact which, perhaps contrary to first glance, is not completely trivial.

Proposition 3.3. *Assume that we are given an orthogonal decomposition $V = \underset{i\in I}{\perp} V_i$. Let J and K be two disjoint subsets of I. Then the submodule $V_J = \underset{i\in J}{\perp} V_i$ of V is disjointly orthogonal to $V_K = \underset{i\in K}{\perp} V_i$, and thus*

$$V_{J\cup K} = V_J \perp V_K.$$

Proof. It follows from Lemma 1.4 above that for any three different indices i, j, k the form q is quasilinear on $V_i \times (V_j + V_k)$, and thus V_i is orthogonal to $V_j \perp V_k$. By iteration, we see that the claim holds if J and K are finite. In the general case, let $x \in V_J$ and $y \in V_K$. There exist finite subsets J', K' of J and K with $x \in V_{J'}$, $y \in V_{K'}$, and thus $q(x + y) = q(x) + q(y)$. This proves that V_J is orthogonal to V_K. □

In the rest of this section, *we assume that V has unique base.*

Definition 3.4. *We call a submodule W of V* basic, *if W is spanned by $\mathfrak{B}_W := \mathfrak{B} \cap W$, and thus W is free with base $\mathfrak{B}_W$. Note that then we have a unique direct decomposition $V = W \oplus U$, where the submodule U is basic with base $\mathfrak{B} \setminus \mathfrak{B}_W$. W and U again are R-modules with unique base. We call U the* complement of W in V, *and write $U = W^c$.*

The theory of basic submodules of V is of utmost simplicity. All of the following is obvious.

Scholium 3.5.

(a) *We have a bijection $W \mapsto \mathfrak{B}_W := \mathfrak{B} \cap W$ from the set of basic submodules of V onto the set of subsets of $\mathfrak{B}$.*

(b) *If W_1 and W_2 are basic submodules of V, then also $W_1 \cap W_2$ and $W_1 + W_2$ are basic submodules of V, and*

$$\mathfrak{B}_{W_1 \cap W_2} = \mathfrak{B}_{W_1} \cap \mathfrak{B}_{W_2}, \qquad \mathfrak{B}_{W_1 + W_2} = \mathfrak{B}_{W_1} \cup \mathfrak{B}_{W_2}.$$

(c) *If W is a basic submodule of V, then as stated above,*

$$\mathfrak{B}_{W^c} = \mathfrak{B} \setminus \mathfrak{B}_W.$$

(d) *Finally, if $W_1 \subset W_2$ are basic submodules of V, then W_1 is basic in W_2 and $W_1^c \cap W_2$ is the complement of W_1 in W_2.*

Thus a basic orthogonal summand W of V has only one basic orthogonal complement, namely, W^c, equipped with the form $q|W^c$.

Definition 3.6. *If the quadratic module V has a basic orthogonal summand $W \neq V$, we call V* decomposable. *Otherwise we call V* indecomposable. *More generally, we call a basic submodule X of V* decomposable *if X is decomposable with respect to $q|X$, and otherwise we call X* indecomposable.

Our next goal is to decompose the given quadratic module V orthogonally into indecomposable basic submodules. Therefore, we choose a base $\mathfrak{B}$ of V (unique up to multiplication by scalar units). We then choose a companion b of q such that $b(\varepsilon, \eta) = 0$ for any two *different* $\varepsilon, \eta \in \mathfrak{B}$ such that q is quasilinear on $R\varepsilon \times R\eta$, cf. [22, Theorem 6.3]. We call such a companion b a *quasiminimal companion* of q.

Comment. In important cases, e.g., if R is supertropical or more generally "upper bound" (cf. [22, Definition 5.1]), the set of companions of q can be partially ordered in a natural way. The prefix "quasi" here is a reminder that we do not mean minimality with respect to such an ordering.

Lemma 3.7. *Let W and W' be basic submodules of V with $W \cap W' = \{0\}$. If b is any quasiminimal companion of q, then W is (disjointly) orthogonal to W' iff $b(W, W') = 0$.*

Proof. If $b(W, W') = 0$, then $q(x+y) = q(x) + q(y)$ for any $x \in W$ and $y \in W'$, which means by definition that W is orthogonal to W'. (This holds for any companion b of q.)
Conversely, if W is orthogonal to W', then for base vectors $\varepsilon \in \mathfrak{B}_W$, $\eta \in \mathfrak{B}_{W'}$ the form q is quasilinear on $R\varepsilon \times R\eta$ and thus $b(\varepsilon, \eta) = 0$. This implies that $b(W, W') = 0$. $\square$

We now introduce the following equivalence relation on the set $\mathfrak{B}$. We choose a quasiminimal companion b of q. Given $\varepsilon, \eta \in \mathfrak{B}$, we put $\varepsilon \sim \eta$, iff either $\varepsilon = \eta$, or there exists a sequence $\varepsilon_0, \varepsilon_1, \dots, \varepsilon_r$ in $\mathfrak{B}$, $r \geq 1$, such that $\varepsilon = \varepsilon_0$, $\eta = \varepsilon_r$, and $\varepsilon_i \neq \varepsilon_{i+1}$, $b(\varepsilon_i, \varepsilon_{i+1}) \neq 0$ for $i = 0, \dots, r-1$.

Theorem 3.8. *Let $\{\mathfrak{B}_k \mid k \in K\}$ denote the set of equivalence classes in $\mathfrak{B}$ and, for every $k \in K$, let W_k denote the submodule of V having base $\mathfrak{B}_k$.*

(a) *Then every* W_k *is an indecomposable basic submodule of* V *and*

$$V = \underset{k \in K}{\perp} W_k.$$

(b) *Every indecomposable basic submodule* U *of* V *is contained in* W_k*, for some* $k \in K$ *uniquely determined by* U.

(c) *The modules* W_k*,* $k \in K$*, are precisely all the indecomposable basic orthogonal summands of* V.

Proof. (a): Suppose that W_k has an orthogonal decomposition $W_k = X \perp Y$ with basic submodules $X \neq 0$, $Y \neq 0$. Then $\mathfrak{B}_k$ is the disjoint union of the non-empty sets $\mathfrak{B}_X$ and $\mathfrak{B}_Y$. Choosing $\varepsilon \in \mathfrak{B}_X$ and $\eta \in \mathfrak{B}_Y$, there exists a sequence $\varepsilon_0, \varepsilon_1, \dots, \varepsilon_r$ in $\mathfrak{B}_k$ with $\varepsilon = \varepsilon_0$, $\eta = \varepsilon_r$ and $b(\varepsilon_{i-1}, \varepsilon_i) \neq 0$, $\varepsilon_{i-1} \neq \varepsilon_i$, for $1 \leq i \leq r$. Let s denote the last index in $\{1, \dots, r\}$ with $\varepsilon_s \in \mathfrak{B}_X$. Then $s < r$ and $\varepsilon_{s+1} \in \mathfrak{B}_Y$. But $b(X, Y) = 0$ by Lemma 3.7 and thus $b(\varepsilon_s, \varepsilon_{s+1}) = 0$, a contradiction. This proves that W_k is indecomposable. Since $\mathfrak{B}$ is the disjoint union of the sets $\mathfrak{B}_k$, we have

$$V = \bigoplus_{k \in K} W_k.$$

Finally, if $k \neq \ell$, then $b(W_k, W_\ell) = 0$ by the nature of our equivalence relation. Thus

$$V = \underset{k \in K}{\perp} W_k.$$

(b): Given an indecomposable basic submodule U of V, we choose $k \in K$ with $\mathfrak{B}_U \cap \mathfrak{B}_k \neq \emptyset$. Then $U \cap W_k \neq 0$. ¿From $V = W_k \oplus W_k^c$, we conclude that $U = (U \cap W_k) \oplus (U \cap W_k^c)$, and then have $U = (U \cap W_k) \perp (U \cap W_k^c)$ because W_k is orthogonal to W_k^c. Since U is indecomposable and $U \cap W_k \neq 0$, it follows that $U = U \cap W_k$, i.e., $U \subset W_k$. Since $W_k \cap W_\ell = 0$ for $k \neq \ell$, it is clear that k is uniquely determined by U.

(c): If U is an indecomposable basic orthogonal summand of V, then $V = U \perp U^c$. We have $U \subset W_k$ for some $k \in K$, and obtain $W_k = U \perp (U^c \cap W_k)$, whence $W_k = U$. □

Definition 3.9. *We call the submodules* W_k *of* V *occurring in Theorem 3.8 the* indecomposable components *of the quadratic module* V.

The following facts are easy consequences of the theorem.

Remark 3.10.

(i) *If* U *is a basic orthogonal summand of* V*, then the indecomposable components of the quadratic module* $U = (U, q|U)$ *are the indecomposable components of* V *contained in* U.

(ii) *If* U *is any basic submodule of* V*, then*

$$U = \underset{k \in K}{\perp} (U \cap W_k),$$

and every submodule $U \cap W_k \neq \{0\}$ is an orthogonal sum of indecomposable components of U.

4. Orthogonal decomposition of bilinear modules with unique base

We now outline a theory of symmetric bilinear forms analogous to the theory for quadratic forms given in §3. The bilinear theory is easier than the quadratic theory due the fact that, in contrast to quadratic forms, on a free module we do not need to distinguish between "functional" and "formal" bilinear forms cf. [22, §1]. As before, R is a semiring.
Assume in the following that V is an R-module equipped with a fixed symmetric bilinear form $b : V \times V \to R$. We then call $V = (V, b)$ a *bilinear R-module.* If X is a submodule of V, we denote the restriction of b to $X \times X$ by $b|X$.

Definition 4.1.

(a) *Given two submodules W_1, W_2 of the R-module V, we say that W_1 is* disjointly orthogonal *to W_2, if $W_1 \cap W_2 = \{0\}$ and $b(W_1, W_2) = 0$, i.e., $b(x, y) = 0$ for all $x \in W_1$, $y \in W_2$.*

(b) *We write $V = W_1 \perp W_2$ if W_1 is disjointly orthogonal to W_2 and moreover $V = W_1 \oplus W_2$ (as R-module). We then call W_1 an* orthogonal summand *of V and W_2 an* orthogonal complement *of W_1 in V.*

Definition 4.2. *Let $(V_i \mid i \in I)$ be a family of submodules of the bilinear module V. We say that V is* the orthogonal sum of the family (V_i), *and then write*

$$V = \mathop{\perp}_{i \in I} V_i,$$

if for any two different indices i, j the submodule V_i is disjointly orthogonal to V_j, and moreover $V = \bigoplus_{i \in I} V_i$.

In contrast to the quadratic case, the exact analog of Proposition 3.3 is now a triviality.

Proposition 4.3. *Assume that $V = \mathop{\perp}_{i \in I} V_i$. Let J and K be disjoint subsets of I. Then $V_J = \mathop{\perp}_{i \in J} V_i$ is disjointly orthogonal to $V_K = \mathop{\perp}_{i \in K} V_i$, and*

$$V_{J \cup K} = V_J \perp V_K.$$

In the following, *we assume again that V has unique base.* Then again a basic orthogonal summand W of V has only one basic orthogonal complement in V, namely, W^c equipped with the bilinear form $b|W^c$.
For X a basic submodule of V, we define the properties *"decomposable"* and *"indecomposable"* in exactly the same way as indicated by Definition 3.6 in the quadratic case.
We start with a definition and description of the "indecomposable components" of $V = (V, b)$ in a similar fashion as was done in §3 for quadratic modules. We choose a base $\mathfrak{B}$ of V and again introduce the appropriate equivalence relation

on the set $\mathfrak{B}$, but now we adopt a more elaborate terminology than in §3. This will turn out to be useful later on.

Definition 4.4. *We call the symmetric bilinear form b* alternate *if $b(\varepsilon, \varepsilon) = 0$ for every $\varepsilon \in \mathfrak{B}$.*

Comment. Beware that this does *not* imply that $b(x, x) = 0$ for every $x \in V$. The classical notion of an alternating bilinear form is of no use here since in the semirings under consideration here (cf. §2) $\alpha + \beta = 0$ implies $\alpha = \beta = 0$, whence $b(x + y, x + y) = 0$ implies $b(x, y) = 0$. An alternating bilinear form in the classical sense would be identically zero.

Definition 4.5. *We associate to the given symmetric bilinear form b an alternate bilinear form b_{alt} by the rule*

$$b_{\mathrm{alt}}(\varepsilon, \eta) = \begin{cases} b(\varepsilon, \eta) & \text{if } \varepsilon \neq \eta \\ 0 & \text{if } \varepsilon = \eta \end{cases}$$

for any $\varepsilon, \eta \in \mathfrak{B}$.

Lemma 4.6. *Let W and W' be basic submodules of V with $W \cap W' = \{0\}$. Then W is (disjointly) orthogonal to W' iff $b_{\mathrm{alt}}(W, W') = 0$.*

Proof. This can be seen exactly as with the parallel Lemma 3.7. Just replace in its proof the quasiminimal companion of q by b_{alt}. □

Definition 4.7.

(a) *A* path Γ *in $V = (V, b)$ of* length $r \geq 1$ *in $\mathfrak{B}$ is a sequence $\varepsilon_0, \varepsilon_1, \dots, \varepsilon_r$ of elements of $\mathfrak{B}$ with*

$$b_{\mathrm{alt}}(\varepsilon_i, \varepsilon_{i+1}) \neq 0 \qquad (0 \leq i \leq r - 1).$$

In essence this condition does not depend on the choice of the base $\mathfrak{B}$, since $\mathfrak{B}$ is unique up to multiplication by units, and so we also say that Γ is a path in V. *We say that the path* runs *from $\varepsilon := \varepsilon_0$ to $\eta := \varepsilon_r$, or that the path* connects *ε to η. A path of length 1 is called an* edge. *This is just a pair (ε, η) in $\mathfrak{B}$ with $\varepsilon \neq \eta$ and $b(\varepsilon, \eta) \neq 0$.*

(b) *We define an equivalence relation on $\mathfrak{B}$ as follows. Given $\varepsilon, \eta \in \mathfrak{B}$, we declare that $\varepsilon \sim \eta$ if either $\varepsilon = \eta$ or there runs a path from ε to η.*

It is now obvious how to mimic the theory of indecomposable components from the end of §3 in the bilinear setting.

Scholium 4.8. *Theorem 3.8 and its proof remain valid for the present equivalence relation on $\mathfrak{B}$. We only have to replace the quasiminimal companion b of q there by b_{alt} and to use Lemma 4.6 instead of Lemma 3.7. Again we denote the set of equivalence classes of $\mathfrak{B}$ by $\{\mathfrak{B}_k \mid k \in K\}$ and the submodule of V with base $\mathfrak{B}_k$ by V_k, and again we call the V_k the* indecomposable components *of V. Also the analog to Remark 3.10 remains valid.*

We state a consequence of the parallel between the two decomposition theories.

Theorem 4.9. *Assume that* (V, q) *is a quadratic module with unique base and* b *is a quasiminimal companion of* q*. The indecomposable components of* (V, q) *coincide with the indecomposable components of* (V, b)*.*

Proof. The equivalence relation used in Theorem 3.8 is the same as the equivalence relation in Definition 4.7. □

We add an easy observation on bilinear modules.

Proposition 4.10. *Assume that* (V, b) *is a bilinear* R*-module with unique base. A basic submodule* W *of* V *is indecomposable with respect to* b*, iff* W *is indecomposable with respect to* b_{alt}*.*

Proof. The equivalence relation on $\mathfrak{B}$ just defined (Definition 4.7) does not change if we replace b by b_{alt}. □

5. Isometries, isotypical components, and a cancelation theorem

Let R be any semiring.

Definition 5.1.

(a) *For quadratic* R*-modules* $V = (V, q)$ *and* $V' = (V', q')$*, an* isometry $\sigma : V \to V'$ *is a bijective* R*-linear map with* $q'(\sigma x) = q(x)$ *for all* $x \in V$*. Likewise, if* $V = (V, b)$ *and* (V', b') *are bilinear* R*-modules, an* isometry *is a bijective* R*-linear map* $\sigma : V \to V'$ *with* $b'(\sigma x, \sigma y) = b(x, y)$ *for all* $x, y \in V$*.*

(b) *If there exists an isometry* $\sigma : V \to V'$*, we call* V *and* V' isometric *and write* $V \cong V'$*. We then also say that* V *and* V' *are in the same isometry class.*

In the following we study quadratic and bilinear R-modules with unique base on an equal footing.

It would not hurt if we supposed that the semiring R satisfies the conditions in Theorem 2.5, so that every free R-module has unique base, but the simplicity of all of the arguments in the present section becomes more apparent if we do not rely on Theorem 2.5.

Notation/Definition 5.2.

(a) *Let* $(V_\lambda^0 \mid \lambda \in \Lambda)$ *be a set of representatives of all isometry classes of indecomposable quadratic (resp. bilinear)* R*-modules with unique base of rank bounded by the cardinality of* V*, in order to avoid set-theoretical complications.*

(b) *If* W *is such an* R*-module, where* $W \cong V_\lambda^0$ *for a unique* $\lambda \in \Lambda$*, we say that* W has type λ *(or:* W *is* indecomposable of type λ*).*

(c) *We say that a quadratic (resp. bilinear) module* $W \neq 0$ *with unique base is* isotypical *of type* λ*, if every indecomposable component of* V *has type* λ*.*

(d) *Finally, given a quadratic (resp. bilinear) R-module with unique base, we denote the sum of all indecomposable components of V of type λ by V_λ and call the $V_\lambda \neq 0$ the* isotypical components of V.

The following is now obvious from §3 and §4 (cf. Theorem 3.8 and Scholium 4.8).

Proposition 5.3. *If V is a quadratic or bilinear R-module with unique base, then*

$$V = \underset{\lambda \in \Lambda'}{\perp} V_\lambda$$

with $\Lambda' = \{\lambda \in \Lambda \mid V_\lambda \neq 0\}$.

Since our notion of orthogonality for basic submodules of V is encoded in the linear and quadratic, resp. bilinear, structure of V, the following fact also is obvious, but in view of its importance will be dubbed a "theorem".

Theorem 5.4. *Assume that V and V' are quadratic (resp. bilinear) R-modules with unique bases and $\sigma : V \to V'$ is an isometry. Let $\{V_k \mid k \in K\}$ denote the set of indecomposable components of V.*

(a) *$\{\sigma(V_k) \mid k \in K\}$ is the set of indecomposable components of V'.*

(b) *If V_k has type λ, then $\sigma(V_k)$ has type λ, and so $\sigma(V_\lambda) = V'_\lambda$ for every $\lambda \in \Lambda$.*

Also in the remainder of the section, we assume that the quadratic or bilinear modules have unique base.

Definition 5.5. *Let $O(V)$ denote the group of all isometries $\sigma : V \to V$ (i.e., automorphisms) of (V, q), resp. (V, b). As usual, we call $O(V)$ the* orthogonal group *of V.*

Theorem 5.4 has the following immediate consequence.

Corollary 5.6. *Every $\sigma \in O(V)$ permutes the indecomposable components of V of fixed type λ, and so $\sigma(V_\lambda) = V_\lambda$ for every $\lambda \in \Lambda$.*
We have a natural isomorphism

$$O(V) \xrightarrow{1:1} \prod_{\lambda \in \Lambda'} O(V_\lambda),$$

sending $\sigma \in O(V)$ to the family of its restrictions $\sigma|V_\lambda \in O(V_\lambda)$.

Definition 5.7.

(a) *Let $\lambda \in \Lambda$. We denote the cardinality of the set of indecomposable components of V_λ by $m_\lambda(V)$, and we call $m_\lambda(V)$ the* multiplicity *of V_λ. $\{$N.B. $m_\lambda(V)$ can be infinite or zero.$\}$*

(b) *If $m_\lambda \in \mathbb{N}_0$ for every $\lambda \in \Lambda$, we say that V is* isotypically finite.

Theorem 5.8. *If V and V' are quadratic or bilinear R-modules with unique bases, then $V \cong V'$ iff $m_\lambda(V) = m_\lambda(V')$ for every $\lambda \in \Lambda$.*

Proof. This follows from Proposition 5.3 and Theorem 5.4. □

We are ready for a main result of the paper.

Theorem 5.9. *Assume that W_1, W_2, W_1', W_2' are quadratic or bilinear modules with unique base and that W_1 is isotypically finite. Assume furthermore that $W_1 \cong W_1'$ and that $W_1 \perp W_2 \cong W_1' \perp W_2'$. Then $W_2 \cong W_2'$.*

Proof. For every $\lambda \in \Lambda$, clearly $m_\lambda(V) = m_\lambda(W_1) + m_\lambda(W_2)$ and $m_\lambda(V') = m_\lambda(W_1') + m_\lambda(W_2')$. Since $V \cong V'$, the multiplicities $m_\lambda(V)$ and $m_\lambda(V')$ are equal, and since $W_1 \cong W_1'$, the same holds for the multiplicities $m_\lambda(W_1')$. Since $m_\lambda(W_1) = m_\lambda(W_1')$ is finite, it follows that $m_\lambda(W_2) = m_\lambda(W_2')$. By Theorem 5.8 this implies that $W_2 \cong W_2'$. □

Remark 5.10. *If the free R-module W_1 has finite rank, then certainly W_1 is isotypically finite. Thus Theorem 5.9 may be viewed as the analog of Witt's cancellation theorem from 1937* [32] *proved for quadratic forms over fields.*

The assumption of isotypical finiteness in Theorem 5.9 cannot be relaxed. Indeed if $m_\lambda(W_1)$ is infinite for at least one $\lambda \in \Lambda$, then the cancelation law becomes false. This is evident by Theorem 5.8 and the following example.

Example 5.11. *Assume that V is the orthogonal sum of infinitely many copies $V_1, V_2, \dots$ of an indecomposable quadratic or bilinear module V_0 with unique base. Consider the following submodules of V:*

$$\begin{aligned} W_1 &:= \ V_2 \perp V_3 \perp \cdots, & W_2 &:= V_1, \\ W_1' &:= \ V_3 \perp V_4 \perp \cdots, & W_2' &:= V_1 \perp V_2. \end{aligned}$$

Then $W_1 \perp W_2 = V = W_1' \perp W_2'$, and $W_1 \cong W_1'$. But W_2 is not isometric to W_2'.

6. Expansions and tensor products

Let $q : V \to R$ be a quadratic form on an R-module V. We recall from [22, §1] that, when V is free with base $(\varepsilon_i : i \in I)$, then q admits a (not necessarily unique) *balanced companion*, i.e., a companion $b : V \times V \to R$ such that $b(x,x) = 2q(x)$ for all $x \in V$, and that it suffices to know for this that $b(\varepsilon_i, \varepsilon_i) = 2q(\varepsilon_i)$ for all $i \in I$ [22, Proposition 1.7]. Balanced companions are a crucial ingredient in our definition below of a tensor product of a free bilinear module and a free quadratic module. They arise from "expansions" of q, defined as follows, cf. [22, Definition 1.9].

Definition 6.1. *A bilinear form $B : V \times V \to R$ (not necessarily symmetric) is an* expansion *of a balanced pair (q, b) if $B + B^t = b$, i.e.,*

$$B(x,y) + B(y,x) = b(x,y) \tag{6.1}$$

for all $x, y \in V$, and

$$q(x) = B(x,x) \tag{6.2}$$

for all $x \in V$. If only the form q is given and (6.2) *holds, we say that B* is an expansion *of q.*

As stated in the [22, §1], every bilinear form $B : V \times V \to R$ gives us a balanced pair (q, b) via (6.1) and (6.2), and, if the R-module V is free, we obtain all such pairs (q, b) in this way. But we will need a description of *all* expansions of (q, b) in the free case.

Construction 6.2. *Assume that V is a free R-module and $(\varepsilon_i \mid i \in I)$ is a base of V. When (q, b) is a balanced pair on V, we obtain all expansions $B : V \times V \to R$ of (q, b) as follows.*
Let $\alpha_i := q(\varepsilon_i)$, $\beta_{ij} := b(\varepsilon_i, \varepsilon_j)$ for $i, j \in I$. We have $\beta_{ij} = \beta_{ji}$. We choose a total ordering on I and for every $i < j$ two elements $\chi_{ij}, \chi_{ji} \in R$ with

$$\beta_{ij} = \chi_{ij} + \chi_{ji}, \qquad (i < j).$$

We furthermore put

$$\chi_{ii} := \alpha_i,$$

and define B by the rule

$$B(\varepsilon_i, \varepsilon_j) = \chi_{ij}$$

for all $(i, j) \in I \times I$.

In practice one usually chooses $\chi_{ij} = \beta_{ij}$, $\chi_{ji} = 0$ for $i < j$, i.e., takes the unique "triangular" expansion B of (q, b), cf. [22, §1], but now we do not want to depend on the choice of a total ordering of the base $(\varepsilon_i \mid i \in I)$. We used such an ordering above only to ease notation.

Tensor products over semirings in general require the use of congruences [10], but for free modules the basics can be done precisely as over rings, and we leave the formal details to the interested reader. We only state here that, given two free R-modules V_1 and V_2, with bases $\mathfrak{B}_1$ and $\mathfrak{B}_2$, the R-module $V_1 \otimes_R V_2$ "is" the free R-module with base $\mathfrak{B}_1 \otimes \mathfrak{B}_2$, which is a renaming of $\mathfrak{B}_1 \times \mathfrak{B}_2$, writing $\varepsilon \otimes \eta$ for (ε, η) with $\varepsilon \in \mathfrak{B}_1$, $\eta \in \mathfrak{B}_2$. If

$$\mathfrak{B}_1 = \{\varepsilon_i \mid i \in I\}, \qquad \mathfrak{B}_2 = \{\eta_j \mid j \in J\}$$

and $x = \sum_{i \in I} x_i \varepsilon_i \in V_1$ and $y = \sum_{j \in J} y_j \eta_j \in V_2$, we define, as common over rings,

$$x \otimes y := \sum_{(i,j) \in I \times J} x_i y_j (\varepsilon_i \otimes y_j), \tag{6.3}$$

and this vector is independent of the choice of the bases $\mathfrak{B}_1$ and $\mathfrak{B}_2$. If B_1 and B_2 are bilinear forms on V_1 and V_2 respectively, we have a well defined bilinear form on $V_1 \otimes_R V_2$, denoted by $B_1 \otimes B_2$, such that for any $x_i \in V_1$, $y_j \in V_2$ $(i, j \in \{1, 2\})$

$$(B_1 \otimes B_2)(x_1 \otimes x_2, y_1 \otimes y_2) = B_1(x_1, y_1) B_2(x_2, y_2). \tag{6.4}$$

If b_1 and b_2 are symmetric bilinear forms on V_1 and V_2 respectively, then $b_1 \otimes b_2$ is symmetric. Then we call the bilinear module $(V_1 \otimes_R V_2, b_1 \otimes b_2)$ the *tensor product of the bilinear modules* (V_1, b_1) and (V_2, b_2).
We next define the tensor product of a free bilinear and a free quadratic module. The key fact which allows us to do this in a reasonable way is as follows.

Proposition 6.3. *Let $\gamma : U \times U \to R$ be a symmetric bilinear form and (q, b) a balanced quadratic pair on V. Assume that B and B' are two expansions of (q, b). Then the bilinear forms $\gamma \otimes B$ and $\gamma \otimes B'$ on $U \otimes V$ yield the same balanced pair $(\tilde{q}, \tilde{b})$ on $U \otimes V$. We have $\tilde{b} = \gamma \otimes b$, whence for $u_1, u_2 \in U$, $v_1, v_2 \in V$,*

$$\tilde{b}(u_1 \otimes v_1, u_2 \otimes v_2) = \gamma(u_1, u_2)b(v_1, v_2). \tag{6.5}$$

Furthermore, for $u \in U$ and $v \in V$,

$$\tilde{q}(u \otimes v) = \gamma(u, u)q(v). \tag{6.6}$$

Proof. $\gamma \otimes B + (\gamma \otimes B)^t = \gamma \otimes B + \gamma^t \otimes B^t = \gamma \otimes B + \gamma \otimes B^t = \gamma \otimes (B + B^t) = \gamma \otimes b$. Also $\gamma \otimes B' + (\gamma \otimes B')^t = \gamma \otimes b$. Furthermore,

$$\begin{aligned}(\gamma \otimes B)(u \otimes v, u \otimes v) &= \gamma(u, u)B(v, v) \\ &= \gamma(u, u)q(v) \qquad = (\gamma \otimes B')(u \otimes v, u \otimes v)\end{aligned}$$

for any $u \in U$, $v \in V$. Together these equations imply

$$(\gamma \otimes B)(z, z) = (\gamma \otimes B')(z, z)$$

for any $z \in U \otimes V$. □

Definition 6.4. *We call $\tilde{q}$ the* tensor product *of the bilinear form γ and the quadratic form q* with respect to the balanced companion *b of q, and write*

$$\tilde{q} = \gamma \otimes_b q,$$

and we also write $\widetilde{V} = U \otimes_b V$ for the quadratic R-module $\widetilde{V} = (U \otimes V, \tilde{q})$.

Remark 6.5. *If q has only one balanced companion, we may suppress the "b" here, writing $\tilde{q} = \gamma \otimes q$. Cases in which this happens are: q is rigid, V has rank one, R is embeddable in a ring.*

Proposition 6.6. *If $U = (U, \gamma)$ has an orthogonal decomposition $U = \underset{i \in I}{\perp} U_i$, then*

$$U \otimes_b V = \underset{i \in I}{\perp} U_i \otimes_b V.$$

Proof. It is immediate that $(\gamma \otimes b)(U_i \otimes V, U_j \otimes V) = 0$ for $i \neq j$. □

We proceed to explicit examples. For this we need notation from [22, §1] which we recall for the convenience of the reader.

Assume that V is free of finite rank n and $\mathfrak{B}$ is a base of V for which we now choose a total ordering, $\mathfrak{B} = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$. Then we identify a bilinear form B on V with the $(n \times n)$-matrix

$$B = \begin{pmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1n} \\ \beta_{21} & \beta_{22} & & \beta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{n1} & \cdots & & \beta_{nn} \end{pmatrix}, \tag{6.7}$$

where $\beta_{ij} = B(\varepsilon_1, \varepsilon_j)$. In particular, a bilinear R-module (V, β) is denoted by a symmetric $(n \times n)$-matrix, namely its Gram matrix $b = (\beta_{ij})_{1 \le i,j \le n}$, where $\beta_{ij} = \beta_{ji} = b(\varepsilon_i, \varepsilon_j)$.
Given a quadratic module (V, q), we choose a triangular expansion

$$B = \begin{pmatrix} \alpha_1 & \alpha_{12} & \cdots & \alpha_{1n} \\ 0 & \alpha_2 & \cdots & \alpha_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & 0 & \alpha_n \end{pmatrix} \tag{6.8}$$

of q and denote q by the triangular scheme

$$q = \begin{bmatrix} \alpha_1 & \alpha_{12} & \cdots & \alpha_{1n} \\ & \alpha_2 & \cdots & \alpha_{2n} \\ & & \ddots & \vdots \\ & & & \alpha_n \end{bmatrix}, \tag{6.9}$$

so that q is given by the polynomial

$$q(x) = \sum_{i=1}^{n} \alpha_i x_i^2 + \sum_{i<j}^{n} \alpha_{ij} x_i x_j.$$

(Such triangular schemes have already been used in the literature when R is a ring, e.g. [28, I §2].) In the case that q is diagonal, i.e., all α_{ij} with $i < j$ are zero, we usually write instead of (6.8) the single row

$$q = [\alpha_1, \alpha_2, \ldots, \alpha_n]. \tag{6.10}$$

Analogously we use for a diagonal symmetric bilinear form b (i.e., $b(\varepsilon_i, \varepsilon_j) = 0$ for $i \neq j$) the notation

$$b = \langle \beta_{11}, \beta_{22}, \ldots, \beta_{nn} \rangle. \tag{6.11}$$

We note that the quadratic form (6.9) has the balanced companion

$$b = \begin{pmatrix} \alpha_1 & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{12} & \alpha_2 & & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{1n} & \cdots & & \alpha_n \end{pmatrix} \tag{6.12}$$

and (6.10), being diagonal, has the balanced companion

$$b = \langle 2\alpha_1, 2\alpha_2, \ldots, 2\alpha_n \rangle. \tag{6.13}$$

Example 6.7. *If $a_1, \ldots, a_n$, $c \in R$, then*

$$\langle a_1, \ldots a_n \rangle \otimes [c] = [a_1 c, \ldots, a_n c]. \tag{6.14}$$

This is evident from Proposition 6.6 and the rule $\langle a \rangle \otimes [c] = [ac]$ for one-dimensional forms which holds by (6.6). *In particular*

$$[a_1, \ldots, a_n] = \langle a_1, \ldots a_n \rangle \otimes [1]. \tag{6.15}$$

Example 6.8. *(As before, R is any semiring.) Assume that $V = (V, q)$ has dimension n, and take a base $\eta_1, \dots, \eta_n$ of V. Let*

$$(U, \gamma) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

with base $\varepsilon_1, \varepsilon_2$. We choose a balanced companion b of V, written as a symmetric $(n \times n)$-matrix $(b(\eta_i, \eta_j))$. We see by the use of the rules (6.5) *and* (6.6) *that*

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes_b q = \left[\begin{array}{c|c} 0 & b \\ \hline & 0 \end{array}\right] \tag{6.16}$$

written with respect to the base

$$\varepsilon_1 \otimes \eta_1, \dots, \varepsilon_1 \otimes \eta_n, \varepsilon_2 \otimes \eta_1, \dots, \varepsilon_2 \otimes \eta_n.$$

This example illustrates dramatically that in general *the tensor product of γ and q depends on the chosen balanced companion b of q*: tensoring q by $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ produces the symmetric matrix of b.

Remark 6.9. *If γ_1 and γ_2 are bilinear forms on the same free R-module U, then the rules* (6.5) *and* (6.6) *imply for any $\lambda_1, \lambda_2 \in R$ that*

$$(\lambda_1\gamma_1 + \lambda_2\gamma_2) \otimes_b q = \lambda_1(\gamma_1 \otimes_b q) + \lambda_2(\gamma_2 \otimes_b q). \tag{6.17}$$

Example 6.10. *Using* (6.17) *with*

$$\gamma_1 = \langle a_1, a_2 \rangle, \quad \gamma_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \lambda_1 = 1, \quad \lambda_2 = \lambda,$$

we obtain from Proposition 6.6 and Example 6.7 that

$$\begin{pmatrix} a_1 & \lambda \\ \lambda & a_2 \end{pmatrix} \otimes_b q = \left[\begin{array}{c|c} a_1 q & \lambda b \\ \hline & a_2 q \end{array}\right]. \tag{6.18}$$

Example 6.11. *Let*

$$q = \begin{bmatrix} 0 & a_{12} & \cdots & a_{1n} \\ & \ddots & \ddots & \vdots \\ & & & a_{n-1,n} \\ & & & 0 \end{bmatrix}$$

with $a_{ij} \in R$ $(i < j)$. Then q is rigid (cf. [22, Proposition 3.4]*; no assumption on R is needed here). Furthermore, let*

$$\gamma = \begin{pmatrix} \gamma_{11} & \cdots & \gamma_{1m} \\ \vdots & & \vdots \\ \gamma_{m1} & \cdots & \gamma_{mm} \end{pmatrix}$$

with $\gamma_{ij} = \gamma_{ji} \in R$. Then we obtain by the rules (6.5) *and* (6.6) *that*

$$\gamma \otimes q = \begin{array}{|c|c|c|c|} \hline 0 & a_{12}\gamma & \cdots & a_{1n}\gamma \\ \hline & 0 & & a_{2n}\gamma \\ \hline & & \ddots & a_{n-1,n}\gamma \\ \hline & & & 0 \\ \hline \end{array} \tag{6.19}$$

More precisely, if the presentations of q *and* γ *above refer to ordered bases* $(\eta_1, \dots, \eta_n)$ *and* $(\varepsilon_1, \dots, \varepsilon_m)$*, respectively, then* (6.19) *refers to the ordered base*

$$(\varepsilon_1 \otimes \eta_1, \dots, \varepsilon_m \otimes \eta_1, \varepsilon_1 \otimes \eta_2, \dots, \varepsilon_m \otimes \eta_n).$$

We now consider the tensor product $\gamma \otimes [a] = \gamma \otimes_b [a]$, cf. Equation (6.10), where b is the unique balanced companion of $[a]$, (6.13). Our starting point is a definition which makes sense for any semiring R and any R-module U.

Definition 6.12. *Let* $\gamma : U \times U \to R$ *be a symmetric bilinear form. The* norm form *of* γ *is the quadratic form* $n(\gamma) : U \to R$ *with*

$$n(\gamma)(x) := \gamma(x, x)$$

for any $x \in U$.

Remark 6.13. *The norm form* $n(\gamma)$ *has the expansion* $\gamma : U \times U \to R$ *and the associated balanced companion* $\gamma + \gamma^{\mathrm{t}} = 2\gamma$*. The norm forms are precisely all the quadratic forms which admit a symmetric expansion. If* U *has a finite base* $\varepsilon_1, \dots, \varepsilon_n$*, then with respect to this base*

$$n(\gamma) = \begin{bmatrix} \gamma_{11} & 2\gamma_{12} & \cdots & 2\gamma_{1m} \\ & \gamma_{22} & & \\ & & \ddots & \vdots \\ & & & \gamma_{mm} \end{bmatrix}, \tag{6.20}$$

where $\gamma_{ij} := \gamma(\varepsilon_i, \varepsilon_j)$.

Proposition 6.14. *Assume that* $U = (U, \gamma)$ *is a free bilinear* R*-module and* $a \in R$*. Then*

$$U \otimes [a] \cong (U,\ a\, n(\gamma)). \tag{6.21}$$

Proof. We realize the form $[a]$ as a quadratic module (V, q) with $V = R\eta$ free of rank 1 and $q(\eta) = a$. $\{q$ has the unique balanced companion $b : V \times V \to R$, with $b(\eta, \eta) = 2a.\}$ The form $\tilde{q} := \gamma \otimes q = \gamma \otimes_b q$ is given by

$$\tilde{q}(x \otimes \eta) = \gamma(x, x)a = (an(\gamma))(x).$$

The claim is obvious. □

Example 6.15. *Assume that* U *has base* $\varepsilon_1, \dots, \varepsilon_m$*. Let* $\gamma_{ij} := \gamma(\varepsilon_i, \varepsilon_j)$*. Then*

$$\gamma \otimes [a] \cong (a\gamma) \otimes [1],$$

and

$$\gamma \otimes [1] = \begin{bmatrix} \gamma_{11} & 2\gamma_{12} & \cdots & 2\gamma_{1n} \\ & \gamma_{22} & & \\ & & \ddots & \vdots \\ & & & \gamma_{mm} \end{bmatrix}, \tag{6.22}$$

where the right hand side refers to the base $\varepsilon_1 \otimes \eta$, $\varepsilon_2 \otimes \eta$, $\dots$, $\varepsilon_m \otimes \eta$.

At a crucial point in §7, we will need an explicit description of the tensor products $\gamma \otimes_b q$ with q indecomposable of rank 2. We start with a general fact.

Proposition 6.16. *Assume that γ is a symmetric bilinear form on a free R-module U and q_1, q_2 are quadratic forms on a free R-module V. Let b_1, b_2 be balanced companions of q_1 and q_2, respectively. Let $q := \lambda_1 q_1 + \lambda_2 q_2$ with $\lambda_1, \lambda_2 \in R$. Then $b := \lambda_1 b_1 + \lambda_2 b_2$ is a balanced companion of q, and*

$$\gamma \otimes_b q = \lambda_1(\gamma \otimes_{b_1} q_1) + \lambda_2(\gamma \otimes_{b_2} q_2). \tag{6.23}$$

This form has the balanced companion $\gamma \otimes b$ (as we know) and

$$\gamma \otimes b = \lambda_1(\gamma \otimes b_1) + \lambda_2(\gamma \otimes b_2). \tag{6.24}$$

Proof. An easy check by use of (6.5) and (6.6). □

Example 6.17. *We take a free module V with base η_1, η_2, and choose with respect to this base*

$$q_1 = \begin{bmatrix} a_1 & 0 \\ & a_2 \end{bmatrix} = [a_1, a_2], \qquad q_2 = \begin{bmatrix} 0 & c \\ & 0 \end{bmatrix}$$

with $a_1, a_2, c \in R$, $c \neq 0$, and the balanced companions

$$b_1 = \begin{pmatrix} 2a_1 & 0 \\ 0 & 2a_2 \end{pmatrix}, \qquad b_2 = \begin{pmatrix} 0 & c \\ c & 0 \end{pmatrix}.$$

Then

$$q := q_1 + q_2 = \begin{bmatrix} a_1 & c \\ & a_2 \end{bmatrix}$$

has the balanced companion

$$b := b_1 + b_2 = \begin{pmatrix} 2a_1 & c \\ c & 2a_2 \end{pmatrix}.$$

For

$$\gamma = \begin{pmatrix} \gamma_{11} & \cdots & \gamma_{1m} \\ \vdots & & \vdots \\ \gamma_{m1} & \cdots & \gamma_{mm} \end{pmatrix}$$

on a free module U with to the base $\varepsilon_1, \dots, \varepsilon_m$, we get

$$\gamma \otimes_{b_1} q_1 = \left[\begin{array}{c|c} a_1 n(\gamma) & 0 \\ \hline & a_2 n(\gamma) \end{array}\right], \qquad \gamma \otimes_{b_2} \begin{bmatrix} 0 & c \\ & 0 \end{bmatrix} = \left[\begin{array}{c|c} 0 & c\gamma \\ \hline & 0 \end{array}\right],$$

cf. (6.19), *and finally*

$$\gamma \otimes_b \begin{bmatrix} a_1 & c \\ & a_2 \end{bmatrix} = \left[\begin{array}{c|c} a_1 n(\gamma) & c\gamma \\ \hline & a_2 n(\gamma) \end{array}\right] \tag{6.25}$$

with respect to the base

$$\varepsilon_1 \otimes \eta_1, \dots, \varepsilon_m \otimes \eta_1, \varepsilon_1 \otimes \eta_2, \dots, \varepsilon_m \otimes \eta_2.$$

Remark 6.18. *¿From* (6.25) *and* (6.18), *we obtain the useful formula*

$$\gamma \otimes_b \begin{bmatrix} a_1 & c \\ & a_2 \end{bmatrix} = \begin{pmatrix} a_1 & c \\ c & a_2 \end{pmatrix} \otimes_{2\gamma} n(\gamma) , \tag{6.26}$$

by use of Example 6.10 for the quadratic pair $(n(\gamma), 2\gamma)$.

From now on, we assume that V has unique base. {We do not need that U has unique base.}

Definition 6.19. *We call a companion b of q* faithful *if b is balanced and quasiminimal.*

Proposition 6.20. *Assume that b is a faithful companion of q, and that $V = W_1 \perp W_2$ is an orthogonal decomposition of V. Then, writing $U \otimes_b W_i$ instead of $U \otimes_{(b|W_i)} W_i$, we have*

$$U \otimes_b V = U \otimes_b W_1 \perp U \otimes_b W_2$$

for any bilinear R-module U.

Proof. $b(W_1, W_2) = 0$, since b is quasiminimal. It follows that

$$(\gamma \otimes b)(U \otimes W_1, U \otimes W_2) = 0.$$

Thus, $\tilde{q} = \gamma \otimes_b q$ is quasilinear on $(U \otimes W_1) \times (U \otimes W_2)$. □

Example 6.21. *Our assumption, that b is faithful, is necessary here. If $V = W_1 \perp W_2$, and b is balanced, but $b(W_1, W_2) \neq 0$, then*

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes_b V = \begin{bmatrix} 0 & b \\ & 0 \end{bmatrix}$$

is not the orthogonal sum of

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes_b W_1 \quad \text{and} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes_b W_2.$$

Example 6.22. *Let $q = [a_1, a_2, \dots, a_n]$ be a diagonal quadratic form. The diagonal symmetric bilinear form*

$$b := \langle 2q_1, \dots, 2a_n \rangle$$

is the unique faithful companion of q. For any bilinear R-module (U, γ), we have

$$\gamma \otimes_b q = \gamma \otimes [a_1] \perp \cdots \perp \gamma \otimes [a_n]. \tag{6.27}$$

Concerning the forms $\gamma \otimes [a_i]$, recall Proposition 6.14 and Example 6.15.

7. Indecomposability in tensor products

In this section, we assume for simplicity that $R \setminus \{0\}$ *is an entire semiring lacking zero sums*. So every free R-module has unique base (cf. Theorem 2.3), and R has no zero divisors. We discuss decomposability first in tensor products of (free) bilinear modules, later in tensor products of bilinear modules with quadratic modules.

Let $V_1 = (V_1, b_1)$ and $V_2 = (V_2, b_2)$ be indecomposable free (symmetric) bilinear modules over R, and let $V := V_1 \otimes V_2 = (V_1 \otimes V_2, b)$ with $b := b_1 \otimes b_2$. We take bases $\mathfrak{B}_1$ and $\mathfrak{B}_2$ of the R-modules V_1, V_2 respectively and then have the base

$$\mathfrak{B} = \mathfrak{B}_1 \otimes \mathfrak{B}_2 := \{\varepsilon \otimes \eta \mid \varepsilon \in \mathfrak{B}_1, \eta \in \mathfrak{B}_2\}$$

of V. Our task is to determine the indecomposable components of V. First we discuss the "trivial" cases.

Remark 7.1. *Assume that V_1 has dimension (= rank) one, so $V_1 \cong \langle a \rangle$ with $a \in R$. If $a \neq 0$, then V is clearly indecomposable. If $a = 0$, then $b_1 \otimes b_2 = 0$, whence V is indecomposable only if also* $\dim V_2 = 1$. *Then $V = \langle 0 \rangle$.*

In all the following, we assume that $V_1 \neq \langle 0 \rangle$, $V_2 \neq \langle 0 \rangle$.
We resort to §4 to describe bases of the indecomposable components of $V = (V, b)$ as the classes in

$$\mathfrak{B} = \{\varepsilon \otimes \eta \mid \varepsilon \in \mathfrak{B}_1, \eta \in \mathfrak{B}_2\}$$

of an equivalence relation given by "paths", cf. Definition 4.7. So a path of length $r \geq 1$ in V, i.e., in $\mathfrak{B}$, is a sequence

$$\Gamma = (\varepsilon_0 \otimes \eta_0, \varepsilon_1 \otimes \eta_1, \dots, \varepsilon_r \otimes \eta_r) \tag{7.1}$$

with

$$b_1(\varepsilon_i, \varepsilon_{i+1}) b_2(\eta_i, \eta_{i+1}) \neq 0 \tag{7.2}$$

and

$$\varepsilon_i \neq \varepsilon_{i+1} \qquad \text{or} \qquad \eta_i \neq \eta_{i+1} \tag{7.3}$$

for $0 \leq i \leq r-1$.
Let us *first assume that both b_1 and b_2 are alternate*, whence also $b = b_1 \otimes b_2$ is alternate. Now condition (7.3) is a consequence of (7.2) and thus can be ignored. We read off from (7.2) that

$$\Gamma_1 = (\varepsilon_0, \varepsilon_1, \dots, \varepsilon_r), \quad \Gamma_2 = (\eta_0, \eta_1, \dots, \eta_r) \tag{7.4}$$

are paths in V_1 and V_2 respectively of same length r. Conversely, given such paths Γ_1 and Γ_2, they combine to a path Γ of length r in V, as written in (7.1). {Here we use the assumption that R has no zero divisors.} We write

$$\Gamma = \Gamma_1 \otimes \Gamma_2. \tag{7.5}$$

We will speak of "cycles" in $\mathfrak{B}_1$, $\mathfrak{B}_2$, $\mathfrak{B}$, in the following obvious way:

Definition 7.2. *Let $\mathfrak{C}$ be a base of a free bilinear R-module W.*

(a) *We denote the length of a path Γ in $\mathfrak{C}$ by $\ell(\Gamma)$.*
(b) *A* cycle *Δ in W with base point $\zeta \in \mathfrak{C}$ is a path $(\zeta_0, \zeta_1, \dots, \zeta_r)$ in $\mathfrak{C}$ with $\zeta_0 = \zeta_r = \zeta$. We say that the cycle Δ is* even *(resp.* odd*) if $\ell(\Delta)$ is* even *(resp.* odd*). We say that Δ is a* 2-cycle *if $\ell(\Delta) = 2$, whence $\Delta = (\zeta, \zeta', \zeta)$ with (ζ, ζ') an edge.*

Lemma 7.3. *Let $\varepsilon, \varepsilon' \in \mathfrak{B}_1$ and $\eta, \eta' \in \mathfrak{B}_2$. Let Γ_1 be a path from ε to ε' of length r and Γ_2 a path from η to η' of length s, and assume that $r \equiv s$* (mod 2). *Then $\varepsilon \otimes \eta \sim \varepsilon' \otimes \eta'$.*

Proof. Assume, without loss of generality, that $s \geq r$, whence $s = r + 2t$ with $t \geq 0$. If $t = 0$, then $\Gamma_1 \otimes \Gamma_2$ is a path from $\varepsilon \otimes \eta$ to $\varepsilon' \otimes \eta'$ in V. If $t > 0$, we replace $\Gamma_1 = (\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_r)$ by

$$\widetilde{\Gamma}_1 = (\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_r, \varepsilon_{r-1}, \varepsilon_r, \ldots)$$

adjoining t copies of the 2-cycle $(\varepsilon_r, \varepsilon_{r-1}, \varepsilon_r)$ to Γ_1. Now $\widetilde{\Gamma}_1 \otimes \Gamma_2$ runs from $\varepsilon \otimes \eta$ to $\varepsilon' \otimes \eta'$. □

Theorem 7.4. *Assume that both b_1 and b_2 are alternate (and $V_1 \neq \langle 0 \rangle$, $V_2 \neq \langle 0 \rangle$, as always).*

a) *If V_1 or V_2 contains an odd cycle, then $V_1 \otimes V_2$ is indecomposable.*
b) *Otherwise $V_1 \otimes V_2$ is the orthogonal sum of two indecomposable components.*

Proof. a): We assume that V_1 contains an odd cycle Δ with base point δ. Let $\varepsilon \otimes \eta$ and $\varepsilon' \otimes \eta'$ be different elements of $\mathfrak{B}$. We want to verify that $\varepsilon \otimes \eta \sim \varepsilon' \otimes \eta'$. We choose a path Γ_1 from ε to ε' in V_1 and a path Γ_2 from η to η' in V_2. If $\ell(\Gamma_1) \equiv \ell(\Gamma_2) \pmod 2$, then we know by Lemma 7.3 that $\varepsilon \otimes \eta \sim \varepsilon' \otimes \eta'$. Now assume that $\ell(\Gamma_1)$ and $\ell(\Gamma_2)$ have different parity. We choose a new path $\widetilde{\Gamma}_1$ from ε to ε' as follows: We first take a path H from ε to the base point δ of Δ, then we run through Δ, then we take the path inverse to H (in the obvious sense) from δ to ε, and finally we run through Γ_1. The length $\ell(\widetilde{\Gamma}_1)$ has different parity than $\ell(\Gamma_1)$ and thus the same parity as $\ell(\Gamma_2)$. We conclude again that $\varepsilon \otimes \eta \sim \varepsilon' \otimes \eta'$.

b): Now assume that both V_1 and V_2 contain only even cycles. This means that both in V_1 and V_2 all paths from a fixed start to a fixed end have length of the same parity. Given $\varepsilon \otimes \eta$ and $\varepsilon' \otimes \eta'$ in $\mathfrak{B}$, every path Γ from $\varepsilon \otimes \eta$ to $\varepsilon' \otimes \eta'$ has the shape $\Gamma_1 \otimes \Gamma_2$ with Γ_1 running from ε to ε', Γ_2 running from η to η', and $\ell(\Gamma_1) = \ell(\Gamma_2)$. Thus, if the paths from ε to ε' have length of different parity than those from η to η', then $\varepsilon \otimes \eta$ cannot be connected to $\varepsilon' \otimes \eta'$ by a path. But $\varepsilon \otimes \eta$ can be connected to $\varepsilon' \otimes \eta''$, where η'' arises from η' by adjoining an edge at the endpoint of η'. We fix some $\varepsilon_0 \in \mathfrak{B}_1$, and $\eta_0, \eta_1 \in \mathfrak{B}_2$ with $b_2(\eta_0, \eta_1) = 1$. Then every element of $\mathfrak{B}$ can be connected by a path to $\varepsilon_0 \otimes \eta_0$ or to $\varepsilon_0 \otimes \eta_1$, but not to both. V has exactly two indecomposable components. □

Remark 7.5. *Assume again that b_1 and b_2 are alternate and $\mathfrak{B}_1$ and $\mathfrak{B}_2$ both contain only even cycles. Let $\varepsilon, \varepsilon' \in \mathfrak{B}_1$ and $\eta, \eta' \in \mathfrak{B}_2$, and choose paths Γ_1 from ε to ε' and Γ_2 from η to η'. As the proof of Theorem 7.4.b has shown, $\varepsilon \otimes \eta$ and $\varepsilon' \otimes \eta'$ lie in the same indecomposable component of $V_1 \otimes V_2$ iff $\ell(\Gamma_1)$ and $\ell(\Gamma_2)$ have the same parity.*

There remains the case that b_1 or b_2 is not alternate.

Theorem 7.6. *Assume that b_1 is* not *alternate and – as before – that $V_1 = (V_1, b_1)$ and $V_2 = (V_2, b_2)$ are indecomposable. Then $(V_1 \otimes V_2, b_1 \otimes b_2)$ is indecomposable.*

Proof. Every path in $V := V_1 \otimes V_2$ with respect to $(b_1)_{\mathrm{alt}} \otimes (b_2)_{\mathrm{alt}}$ is also a path with respect to $b_1 \otimes b_2$, as is easily checked, and the paths in V_i with respect to b_i are the same as those with respect to $(b_i)_{\mathrm{alt}}$ $(i = 1, 2)$. Thus we are done by Theorem 7.4, except in the case that all cycles in V_1 and in V_2 are even. Then V has two indecomposable components W', W'' with respect to $(b_1)_{\mathrm{alt}} \otimes (b_2)_{\mathrm{alt}}$. The base

$$\mathfrak{B} = \mathfrak{B}_1 \otimes \mathfrak{B}_2 := (\varepsilon \otimes \eta \mid \varepsilon \in \mathfrak{B}_1, \eta \in \mathfrak{B}_2)$$

of $V_1 \otimes V_2$ is the disjoint union of sets $\mathfrak{B}'$, $\mathfrak{B}''$ which are bases of W' and W''. Any two elements of $\mathfrak{B}'$ are connected by a path with respect to $(b_1)_{\mathrm{alt}} \otimes (b_2)_{\mathrm{alt}}$, hence by a path with respect to $b_1 \otimes b_2$, and the same holds for the set $\mathfrak{B}''$.
We choose some $\rho \in \mathfrak{B}_1$ with $b_1(\rho, \rho) \neq 0$ and an edge (η_0, η_1) in $\mathfrak{B}''$. Since R has no zero divisors, it follows that $(\rho \otimes \eta_0, \rho \otimes \eta_1)$ is an edge in $\mathfrak{B}$ with respect to $b_1 \otimes b_2$. Perhaps interchanging W' and W'', we assume that $\rho \otimes \eta_0 \in \mathfrak{B}'$. Suppose that also $\rho \otimes \eta_1 \in \mathfrak{B}'$. Then there exists a path Γ in $\mathfrak{B}'$ with respect to $(b_1)_{\mathrm{alt}} \otimes (b_2)_{\mathrm{alt}}$ running from $\rho \otimes \eta_0$ to $\rho \otimes \eta_1$. Γ has the form $\Gamma_1 \otimes \Gamma_2$, with Γ_1 a cycle in V_1 with base point ρ, and Γ_2 a path in V_2 running from η_0 to η_1. We have $\ell(\Gamma_1) = \ell(\Gamma_2)$ and $\ell(\Gamma_2)$ is even. But there exists the path (η_0, η_1) from η_0 to η_1 of length 1. Since all paths in V_2 from η_0 to η_1 have the same parity, we infer that $\ell(\Gamma_2)$ is odd, a contradiction.
We conclude that $\rho \otimes \eta_1 \in \mathfrak{B}''$. The elements $\rho \otimes \eta_0 \in \mathfrak{B}'$ and $\rho \otimes \eta_1 \in \mathfrak{B}''$ are connected by a path with respect to $b_1 \otimes b_2$, and thus all elements of $\mathfrak{B}$ are connected by paths with respect to $b_1 \otimes b_2$. □

Turning to a study of indecomposable components of tensor products of bilinear and quadratic modules, we need some more terminology. Let $V = (V, q)$ be a free quadratic R-module and $\mathfrak{B}$ a base of V. We focus on balanced companions of q.

Definition 7.7.

(a) *We call a companion b of q* faithful *if b is* balanced and quasiminimal *(cf. §3 above), whence $b(\varepsilon, \varepsilon) = 2q(\varepsilon)$ for all $\varepsilon \in \mathfrak{B}$ and $b(\varepsilon, \eta) = 0$ for $\varepsilon \neq \eta$ in $\mathfrak{B}$ such that q is quasilinear on $R\varepsilon \times R\eta$.*

(b) *Given a balanced companion b of q, we define a new bilinear form b_f on V by the rule that, for $\varepsilon, \eta \in \mathfrak{B}$,*

$$b_f(\varepsilon, \eta) = \begin{cases} 0 & \text{if } \varepsilon \neq \eta \text{ and } q \text{ is quasilinear on } R\varepsilon \times R\eta, \\ b(\varepsilon, \eta) & \text{else.} \end{cases}$$

It is clear from [22, Theorem 6.3] that again b_f is a companion of q. By definition, this companion is quasiminimal. b_f is also balanced, since $b_f(\varepsilon, \varepsilon) = b(\varepsilon, \varepsilon) = 2q(\varepsilon)$ for all $\varepsilon \in \mathfrak{B}$, cf. [22, Proposition 1.7], and so b_f is faithful. We call b_f the *faithful companion of q associated to b.*

Theorem 7.8. *Assume that b is a balanced companion of q, and that W is a basic submodule of V. Then W is indecomposable with respect to q iff W is indecomposable with respect to b_f.*

Proof. This is a special case of Theorem 4.9, since $b_f|W = (b|W)_f$ is a quasi-minimal companion of $q|W$. $\square$

Definition 7.9.

(a) *We say that q is* diagonally zero *if $q(\varepsilon) = 0$ for every $\varepsilon \in \mathfrak{B}$.*

(b) *We say that q is* anisotropic *if $q(\varepsilon) \neq 0$ for every $\varepsilon \in \mathfrak{B}$.*

Remarks 7.10.

(i) *If q is diagonally zero, then q is rigid, cf.* [22, Proposition 3.4]. *Conversely, if q is rigid and the quadratic form* [1] *is quasilinear, i.e., $(\alpha+\beta)^2 = \alpha^2+\beta^2$ for any $\alpha, \beta \in R$, then q is diagonally zero, as proved in* [22, Theorem 3.5].

(ii) *If q is anisotropic, then $q(x) \neq 0$ for every $x \in V \setminus \{0\}$. So our definition of anisotropy here coincides with the usual meaning of anisotropy for quadratic forms (which makes sense, say, for R a semiring without zero divisors and V any R-module).*

Definition 7.11. *In a similar vein, we call a symmetric* bilinear form *b on V* anisotropic *if $b(\varepsilon, \varepsilon) \neq 0$ for every $\varepsilon \in \mathfrak{B}$, and then have $b(x,x) \neq 0$ for every $x \in V \setminus \{0\}$.*

Note that, if b is a balanced companion of q, then b is anisotropic iff q is anisotropic.

Assume now that $U := (U, \gamma)$ is a free bilinear module, $V := (V, q)$ is a free quadratic module, and b is a balanced companion of q. Let

$$\widetilde{V} := (\widetilde{V}, \tilde{q}) := (U \otimes V, \gamma \otimes_b q).$$

We want to determine the indecomposable components of $\widetilde{V}$. Discarding trivial cases, we assume that $U \neq \langle 0 \rangle$, $V \neq [0]$.

We choose bases $\mathfrak{B}_1$ and $\mathfrak{B}_2$ of the R-modules U and V, respectively, and introduce the subsets

$$\begin{aligned}
\mathfrak{B}_1^+ &:= \{\varepsilon \in \mathfrak{B}_1 \mid \gamma(\varepsilon, \varepsilon) \neq 0\}, \\
\mathfrak{B}_1^0 &:= \{\varepsilon \in \mathfrak{B}_1 \mid \gamma(\varepsilon, \varepsilon) = 0\}, \\
\mathfrak{B}_2^+ &:= \{\eta \in \mathfrak{B}_1 \mid q(\eta) \neq 0\}, \\
\mathfrak{B}_2^0 &:= \{\eta \in \mathfrak{B}_1 \mid q(\eta) = 0\},
\end{aligned}$$

of $\mathfrak{B}_1$ and $\mathfrak{B}_2$, respectively, and furthermore the basic submodules U^+, U^0, V^+, V^0 respectively spanned by these sets.

Lemma 7.12.

a) *If $\varepsilon \in \mathfrak{B}_1^+$, then the indecomposable components of the basic submodule $\varepsilon \otimes V := (R\varepsilon) \otimes V$ of $U \otimes V$ with respect to $\tilde{q}$ are the submodules $\varepsilon \otimes W$ with W running through the indecomposable components of V with respect to q.*

b) *If $\eta \in \mathfrak{B}_1^+$, then the indecomposable components of $U \otimes \eta := U \otimes (R\eta)$ with respect to $\tilde{q}$ are the modules $U \otimes \eta$ with U' running through the indecomposable components of U with respect to the norm form $n(\gamma)$ of γ (cf. Definition 6.12).*

Proof. This follows from the formulas $\tilde{q}(\varepsilon \otimes y) = \gamma(\varepsilon, \varepsilon)q(y)$ for $y \in V$ and $\tilde{q}(x \otimes \eta) = \gamma(x, x)q(\eta)$ for $x \in U$ (cf. (6.6)), since $\gamma(\varepsilon, \varepsilon) \neq 0$, $q(\eta) \neq 0$. □

In order to avoid certain pathologies concerning indecomposability in tensor products $U \otimes_b V$, we henceforth will assume that our semiring has the following property:

(NQL) For any a and c in $R \setminus \{0\}$ there exists some $\mu \in R$ with $a + \mu c \neq a$.

Clearly, this property means that every free quadratic module $\left[\begin{smallmatrix} a & c \\ & 0 \end{smallmatrix}\right]$ with $c \neq 0$ is *not quasilinear* on $(R\eta_1) \times (R\eta_2)$, where (η_1, η_2) is the associated base, whence the label “NQL”.

Examples 7.13.

(a) *In the important case that R is supertropical the condition (*NQL*) holds iff all principal ideals in eR are unbounded with respect to the total ordering of eR. In particular, the “multiplicatively unbounded supertropical semirings” appearing in* [20, §7] *have* NQL.

(b) *If R is any entire semiring lacking zero sums, then the polynomial ring $R[t]$ in one variable (and so in any set of variables) has* NQL.

(c) *The polynomial function semirings over supersemirings appearing in* [25, §4] *have* NQL.

Lemma 7.14. *Assume that (V, q) is indecomposable. Let $a, c \in R \setminus \{0\}$. Then*

$$\begin{pmatrix} a & c \\ c & 0 \end{pmatrix} \otimes_b V = \left[\begin{array}{c|c} aq & cb \\ \hline & 0 \end{array}\right]$$

(cf. (6.19)*) is indecomposable.*

Proof. Let

$$(U, \gamma) = \begin{pmatrix} a & c \\ c & 0 \end{pmatrix}$$

with respect to a base $\varepsilon_1, \varepsilon_2$ and assume for notational convenience that V has a finite base $\eta_1, \dots, \eta_n$. By Lemma 7.12.a, we have

$$\varepsilon_1 \otimes \eta_1 \sim \varepsilon_1 \otimes \eta_2 \sim \cdots \sim \varepsilon_1 \otimes \eta_n.$$

For given $\varepsilon_1 \otimes \eta_i$, $\varepsilon_2 \otimes \eta_j$ with $i \neq j$, $\gamma \otimes_b q$ has the value table

$$\begin{bmatrix} aq(\eta_i) & cb(\eta_i, \eta_j) \\ & 0 \end{bmatrix}.$$

Starting with $\varepsilon_2 \otimes \eta_j$, we find some η_i, $i \neq j$, with $b(\eta_i, \eta_j) \neq 0$, because (V, q) is indecomposable. Since R has NQL, it follows that $R(\varepsilon_i \otimes \eta_i) + R(\varepsilon_j \otimes \eta_j)$ is indecomposable with respect to $\tilde{q}$, whence $\varepsilon_1 \otimes \eta_i \sim \varepsilon_2 \otimes \eta_j$. Thus all $\varepsilon_k \otimes \eta_\ell$ are equivalent. □

Lemma 7.15. *Assume that $(U, n(\gamma))$ is indecomposable. Let $a, c \in R \setminus \{0\}$. Then the tensor product $U \otimes_b \begin{bmatrix} a & c \\ & 0 \end{bmatrix}$, taken with respect to $b = \begin{pmatrix} 2a & c \\ c & 0 \end{pmatrix}$, is indecomposable.*

Proof. By formula (6.26)

$$\gamma \otimes_b \begin{bmatrix} a & c \\ & 0 \end{bmatrix} = \begin{pmatrix} a & c \\ c & 0 \end{pmatrix} \otimes_{2\gamma} n(\gamma).$$

Now Lemma 7.14 with $(V, q) := (U, n(\gamma))$ gives the claim. ☐

We are ready for the main result of this section. Recall that $U := (U, \gamma)$.

Theorem 7.16. *Assume that R has* NQL. *Assume furthermore that both $(U, n(\gamma))$ and the quadratic free module $V = (V, q)$ are indecomposable, and $U \neq \langle 0 \rangle$, $V \neq [0]$. Let b be a balanced companion of q. Then the quadratic module $U \otimes_b V := (U \otimes V, \gamma \otimes_b q)$ is indecomposable, except in the case that γ is alternate, q is diagonally zero, U and V contain only even cycles with respect to γ and b. Then $U \otimes_b V$ has exactly two indecomposable components, and these coincide with the indecomposable components of $U \otimes V$ with respect to $\gamma \otimes b$, and also with respect to $\gamma \otimes b_f$.*

Proof. Of course, indecomposability of $(U, n(\gamma))$ implies indecomposability of (U, γ). As before, let $\tilde{q} := \gamma \otimes_b q$. We distinguish three cases.

1) Assume that $V^+ \neq \{0\}$, i.e., there exist anisotropic base vectors in V. Our claim is that all elements of $\mathfrak{B}_1 \otimes \mathfrak{B}_2$ are equivalent, whence $U \otimes_b V$ is indecomposable.
We choose $\eta_0 \in \mathfrak{B}_2^+$. By Lemma 7.12.b, the module

$$(U \otimes \eta_0, \tilde{q}) := (U \otimes \eta_0, \tilde{q} \mid U \otimes \eta_0)$$

is indecomposable, and thus all elements of $\mathfrak{B}_1 \otimes \eta_0$ are equivalent.
Let $\varepsilon \otimes \eta \in \mathfrak{B}_1 \otimes \mathfrak{B}_2$. We verify the equivalence of $\varepsilon \otimes \eta$ with some element of $\mathfrak{B}_1 \otimes \eta_0$, and then will be done. If $\gamma(\varepsilon, \varepsilon) \neq 0$, then by Lemma 7.12.a, all elements of $\varepsilon \otimes \mathfrak{B}_2$ are equivalent, whence $\varepsilon \otimes \eta \sim \varepsilon \otimes \eta_0$. Assume now that $\gamma(\varepsilon, \varepsilon) = 0$. Since (U, γ) is indecomposable, there exists some $\varepsilon' \in \mathfrak{B}_1$ with $c := \gamma(\varepsilon', \varepsilon) \neq 0$. Let $a := \gamma(\varepsilon', \varepsilon')$. We choose a base $\eta_1, \dots, \eta_n$ of V, assuming for notational convenience that V has finite rank. By Example 6.10,

$$(R\varepsilon' + R\varepsilon) \otimes_b V = \begin{bmatrix} aq & cq \\ & 0 \end{bmatrix}$$

with respect to the base $\varepsilon' \otimes \eta_1, \dots, \varepsilon' \otimes \eta_n, \varepsilon \otimes \eta_1, \dots, \varepsilon \otimes \eta_2$. Now Lemma 7.14 tells us that $(R\varepsilon' + R\varepsilon) \otimes_b V$ is indecomposable, whence all elements $\varepsilon \otimes \eta$, $\varepsilon' \otimes \eta'$ with $\eta, \eta' \in \mathfrak{B}_2$ are equivalent. In particular, $\varepsilon \otimes \eta \sim \varepsilon' \otimes \eta_0$.

2) Assume that $U^+ \neq \{0\}$, i.e., there exist an anisotropic base vector in U with respect to $n(\gamma)$. Our claim again is that all elements of $\mathfrak{B}_1 \otimes \mathfrak{B}_2$ are equivalent, whence $U \otimes_b V$ is indecomposable. We choose $\varepsilon_0 \in \mathfrak{B}_1^+$, and then know by Lemma 7.12.a that all elements of $\varepsilon_0 \otimes \mathfrak{B}_2$ are equivalent.

Let $\varepsilon \otimes \eta \in \mathfrak{B}_1 \otimes \mathfrak{B}_2$ be given. We verify equivalence of $\varepsilon \otimes \eta$ with some element of $\varepsilon_0 \otimes \mathfrak{B}_2$, and then will be done. If $q(\eta) \neq 0$, then by Lemma 7.12.a all elements of $\mathfrak{B}_1 \otimes \eta$ are equivalent, and thus $\varepsilon \otimes \eta \sim \varepsilon_0 \otimes \eta$.
Hence, we may assume that $q(\eta) = 0$. Since (V, q) is indecomposable, there exists some $\eta' \in \mathfrak{B}_2$ with $c := b(\eta, \eta') \neq 0$. Let $a := q(\eta')$. Then

$$(R\eta' + R\eta, q) = \begin{bmatrix} a & c \\ & 0 \end{bmatrix}.$$

Let $b' := b|(R\eta' + R\eta) = \begin{pmatrix} a & c \\ c & 0 \end{pmatrix}$. Then we see from (6.25) that

$$\gamma \otimes_{b'} \begin{bmatrix} a & c \\ & 0 \end{bmatrix} = \left[\begin{array}{c|c} an(\gamma) & c\gamma \\ \hline & 0 \end{array}\right].$$

By Lemma 7.15, this quadratic module is indecomposable, whence all elements $\varepsilon \otimes \eta$, $\varepsilon' \otimes \eta'$ with $\varepsilon, \varepsilon' \in \mathfrak{B}_1$ are equivalent. In particular, $\varepsilon \otimes \eta \sim \varepsilon_0 \otimes \eta'$.

3) The remaining case: $U = U^0$, and $V = V^0$, i.e., γ is alternate and q is diagonally zero. Now $(U \otimes V, \tilde{q})$ is rigid. By Theorem 7.8, the indecomposable components of $(U \otimes V, \tilde{q})$ coincide with those of $(U \otimes V, (\gamma \otimes b)_f)$. But $\tilde{q}$ has only one companion, whence $(\gamma \otimes b)_f = \gamma \otimes b = \gamma \otimes b_f$. Invoking Theorem 7.4, we see that the assertion of the theorem also holds in the case under consideration, where γ is alternate and b is diagonally zero. □

In general, let $\{U_i \mid i \in I\}$ denote the set of indecomposable components of $(U, n(\gamma))$. Then

$$U \otimes_b V = \perp_{i \in I} U_i \otimes_b V$$

by Proposition 6.6, whence, applying Theorem 7.16 to each summand $U_i \otimes_b V$, we obtain a complete list of all indecomposable components of $U \otimes_b V$. In particular, if q is not diagonally zero, or if (V, b) contains an odd cycle, then the $U_i \otimes_b V$ themselves are the indecomposable components of $U \otimes_b V$.

References

[1] Z. Arad and H.I. Blau. On table algebras and applications to finite group theory, *J. of Alg.* 138:137–185, 1991.

[2] H.I. Blau. Table algebras, *European J. Combin.*, 30(6):1426–1455, 2009

[3] N. Bourbaki. *Alg. Comm. VI*, §3, No.1.

[4] J. Bochnak, M. Coste, and M.-F. Roy. *Real algebraic geometry*, Ergebnisse Math. Grenzgeb. 36, 1998.

[5] G.W. Brumfiel. *Partially ordered rings and semialgebraic geometry*, London Math. Soc. Lecture Notes 37, Cambridge Univ. Press, 1979.

[6] D. Dolžna and P. Oblak. Invertible and nilpotent matrices over antirings. *Lin. Alg. and Appl.*, 430(1):271–278, 2009.

[7] N. Durov. A new approach to Arakelov geometry, arXiv 0704.2030, 2007.

[8] R. Elman, N. Karpenko, and A. Merkurjev. *The Algebraic and Geometric Theory of Quadratic Forms*, Amer. Math. Soc. Colloquium Publications 56, Amer. Math. Soc., Providence, RI, 2008.

[9] J. Golan, *Semirings and their Applications*, Springer-Science + Business, Dordrecht, 1999. (Previously published by Kluwer Acad. Publ., 1999.)

[10] R. Harari. *Noncommutative Supertropical Semirings and Symbol Semialgebras*, Doctoral dissertation, Bar-Ilan University 2015.

[11] D.K. Harrison and M.A. Vitulli. V-valuations of a commutative ring I, *J. Algebra*, 126:264–292, 1989.

[12] D.K. Harrison and M.A. Vitulli. Complex-valued places and CMC subsets of a field, *Commun. in Algebra* 17:2529–2537, 1989.

[13] G. Hoheisel, Über Charaktere, *Monatsh. Math.* 48:448–456, 1939.

[14] R. Huber and M. Knebusch. On valuation spectra, *Contemp. Math.*, 155:167–206, 1994.

[15] I. Itenberg, G. Mikhalkin, and E. Shustin. *Tropical algebraic geometry*, Oberwolfach Seminars, vol. 35, Birkhauser, 2007.

[16] Z. Izhakian. Tropical arithmetic and matrix algebra, *Commun. in Alg.*, 37(4):1445–1468, 2009.

[17] Z. Izhakian, M. Knebusch, and L. Rowen. Supertropical semirings and supervaluations. *J. of Pure and Appl. Alg.*, 215(10):2431–2463, 2011.

[18] Z. Izhakian, M. Knebusch, and L. Rowen. Dominance and transmissions in supertropical valuation theory, *Commun. in Algebra*, 41(7):2736–2782, 2013.

[19] Z. Izhakian, M. Knebusch, and L. Rowen. Monoid valuations and value ordered supervaluations, *Commun. in Alg.*, 43(8):3207-3248, 2015.

[20] Z. Izhakian, M. Knebusch, and L. Rowen. Minimal orderings and quadratic forms on a free module over a supertropical semiring, *Lin. Alg. and Appl.*, 507:420-461, 2016.

[21] Z. Izhakian, M. Knebusch, and L. Rowen. Supertropical linear algebra, *Pacific J. of Math.*, 266(1):43–75, 2013.

[22] Z. Izhakian, M. Knebusch, and L. Rowen. Supertropical quadratic forms I, *J. of Pure and Appl. Alg.*, 220(1):61–93, 2016. (Preprint at arXiv:1309.5729.v2, 2013.)

[23] Z. Izhakian, M. Knebusch, and L. Rowen. Layered tropical mathematics, *J. of Alg.*, 416:200–273, 2014.

[24] Z. Izhakian, M. Knebusch, and L. Rowen. *Supertropical Quadratic Forms II*, arXiv:1506.03404, 2015.

[25] Z. Izhakian and L. Rowen. Supertropical algebra. *Adv. in Math.*, 225(4):2222–2286, 2010.

[26] M. Knebusch. *Specialization of Quadratic and Symmetric Bilinear Forms*, Springer London, 2010.

[27] M. Knebusch and D. Zhang. *Manis Valuations and Prüfer Extensions. I. A New Chapter in Commutative Algebra*, Lecture Notes in Mathematics, 1791, Springer-Verlag, Berlin, 2002.

[28] M. Kneser. *Quadratische Formen*, Springer-Verlag, Berlin-Heidelberg, 2002.

[29] D. Maclagan and B. Sturmfels, *Introduction to Tropical Geometry*, American Mathematical Society, 2015.

[30] Y. Tan. On invertible matrices over antirings. *Lin. Alg. and Appl.*, 423(2):428–444, 2007.
[31] K.G. Valente and M.A. Vitulli. Complex-valued places and CMC subsets of a ring, *Commun. in Algebra*, 18:3743–3757, 1990.
[32] E. Witt. Theorie der quadratischen Formen in beliebigen Körpen. *J. Reine Angew. Math.*, 176:31–44, 1937.
[33] D. Zhang. The m-valuation spectrum of a commutative ring, *Commun. in Algebra*, 30:2883–2896, 2002.

Zur Izhakian
Institute of Mathematics
University
of Aberdeen
AB24 3UE
Aberdeen
UK
zzur@abdn.ac.uk;
zzur@math.biu.ac.il

Manfred Knebusch
Department of Mathematics
NWF-I Mathematik
Universität Regensburg
93040 Regensburg
Germany
manfred.knebusch@
mathematik.uni-
regensburg.de

Louis Rowen
Department of Mathematics
Bar-Ilan
University
Ramat-Gan 52900
Israel
rowen@math.biu.ac.il

C*-Algebras of Boolean Inverse Monoids – Traces and Invariant Means

Charles Starling[1]

Received: May 19, 2016
Revised: June 21, 2016

Communicated by Joachim Cuntz

Abstract. To a Boolean inverse monoid S we associate a universal C*-algebra $C_B^*(S)$ and show that it is equal to Exel's tight C*-algebra of S. We then show that any invariant mean on S (in the sense of Kudryavtseva, Lawson, Lenz and Resende) gives rise to a trace on $C_B^*(S)$, and vice-versa, under a condition on S equivalent to the underlying groupoid being Hausdorff. Under certain mild conditions, the space of traces of $C_B^*(S)$ is shown to be isomorphic to the space of invariant means of S. We then use many known results about traces of C*-algebras to draw conclusions about invariant means on Boolean inverse monoids; in particular we quote a result of Blackadar to show that any metrizable Choquet simplex arises as the space of invariant means for some AF inverse monoid S.

2010 Mathematics Subject Classification: 20M18, 46L55, 46L05

1 Introduction

This article is the continuation of our study of the relationship between inverse semigroups and C*-algebras. An *inverse semigroup* is a semigroup S for which every element $s \in S$ has a unique "inverse" s^* in the sense that

$$ss^*s = s \text{ and } s^*ss^* = s^*.$$

An important subsemigroup of any inverse semigroup is its set of idempotents $E(S) = \{e \in S \mid e^2 = e\} = \{s^*s \mid s \in S\}$. Any set of partial isometries closed under product and involution inside a C*-algebra is an inverse semigroup, and its set of idempotents forms a commuting set of projections. Many C*-algebras A have been profitably studied in the following way:

[1]Supported by the NSERC grants of Benoît Collins, Thierry Giordano, and Vladimir Pestov. `cstar050@uottawa.ca`.

1. identify a generating inverse semigroup S,
2. write down an abstract characterization of S,
3. show that A is universal for some class of representations of S.

We say "some class" above because typically considering all representations (as in the construction of Paterson [Pat99]) gives us a larger C*-algebra than we started with. For example, consider the multiplicative semigroup inside the Cuntz algebra $\mathcal{O}_2$ generated by the two canonical generators s_0 and s_1; in semigroup literature this is usually denoted P_2 and called the *polycyclic monoid* of order 2. The C*-algebra which is universal for all representations of P_2 is $\mathcal{T}_2$, the Toeplitz extension of $\mathcal{O}_2$. In an effort to arrive back at the original C*-algebra in cases such as this, Exel defined the notion of *tight* representations [Exe08], and showed that the universal C*-algebras for tight representations of P_2 is $\mathcal{O}_2$. See [Sta16], [Sta15], [EP16], [EP14], [EGS12], [COP15] for other examples of this approach.
Another approach to this issue is to instead alter the inverse semigroup S. An inverse semigroup carries with it a natural order structure, and when an inverse semigroup S is represented in a C*-algebra A, two elements $s, t \in S$, which did not have a lowest upper bound in S, may have one inside A. So, from P_2, Lawson and Scott [LS14, Proposition 3.32] constructed a new inverse semigroup C_2, called the *Cuntz inverse monoid*, by adding to P_2 all possible joins of *compatible* elements (s, t are compatible if $s^*t, st^* \in E(S)$).
The Cuntz inverse monoid is an example of a *Boolean inverse monoid*, and the goal of this paper is to define universal C*-algebras for such monoids and study them. A Boolean inverse monoid is an inverse semigroup which contains joins of all finite compatible sets of elements and whose idempotent set is a Boolean algebra. To properly represent a Boolean inverse monoid S, one reasons, one should insist that the join of two compatible $s, t \in S$ be sent to the join of the images of s and t. We prove in Proposition 3.3 that such a representation is necessarily a tight representation, and so we obtain that the universal C*-algebra of a Boolean inverse monoid (which we denote $C^*_B(S)$) is exactly its tight C*-algebra, Theorem 3.5. This is the starting point of our study, as the universal tight C*-algebra can be realized as the C*-algebra of an ample groupoid.
The main inspiration of this paper is [KLLR16] which defines and studies *invariant means* on Boolean inverse monoids. An invariant mean is a function $\mu : E(S) \to [0, \infty)$ such that $\mu(e \vee f) = \mu(e) + \mu(f)$ when e and f are orthogonal, and such that $\mu(ss^*) = \mu(s^*s)$ for all $s \in S$. If one thinks of the idempotents as clopen sets in the Stone space of the Boolean algebra $E(S)$, such a function has the flavour of an invariant measure or a trace. We make this precise in Section 4: as long as S satisfies a condition which guarantees that the induced groupoid is Hausdorff (which we call condition (H)), every invariant mean on S gives rise to a trace on $C^*_B(S)$ (Proposition 4.6) and every trace on $C^*_B(S)$ gives rise to an invariant mean on S (Proposition 4.7). This becomes a one-

to-one correspondence if we assume that the associated groupoid $\mathcal{G}_{\text{tight}}(S)$ is principal and amenable (Theorem 4.13). We also prove that, whether $\mathcal{G}_{\text{tight}}(S)$ is principal and amenable or not, there is an affine isomorphism between the space of invariant means on S and the space of $\mathcal{G}_{\text{tight}}(S)$-invariant measures on its unit space (Proposition 4.11).

In the final section, we apply our results to examples of interest. We study the *AF inverse monoids* in detail – these are Boolean inverse monoids arising from Bratteli diagrams in much the same way as AF C*-algebras. As it should be, given a Bratteli diagram, the C*-algebra of its Boolean inverse monoid is isomorphic to the AF algebra it determines (Theorem 5.1). From this we can conclude, using the results of Section 4 and the seminal result of Blackadar [Bla80], that any Choquet simplex arises as the space of invariant means for some Boolean inverse monoid. We go on to consider two examples where there is typically only one invariant mean, those being self-similar groups and aperiodic tilings.

2 Preliminaries and notation

We will use the following general notation. If X is a set and $U \subset X$, let Id_U denote the map from U to U which fixes every point, and let 1_U denote the characteristic function on U, i.e. $1_U : X \to \mathbb{C}$ defined by $1_U(x) = 1$ if $x \in U$ and $1_U(x) = 0$ if $x \notin U$. If F is a finite subset of X, we write $F \subset_{\text{fin}} X$.

2.1 Inverse semigroups

An *inverse semigroup* is a semigroup S such that for all $s \in S$, there is a unique element $s^* \in S$ such that

$$ss^*s = s, \qquad s^*ss^* = s^*.$$

The element s^* is called the *inverse* of s. All inverse semigroups in this paper are assumed to be discrete and countable. For $s, t \in S$, one has $(s^*)^* = s$ and $(st)^* = t^*s^*$. Although not implied by the definition, we will always assume that inverse semigroups have a 0 element, that is, an element such that

$$0s = s0 = 0 \text{ for all } s \in S.$$

An inverse semigroup with identity is called an *inverse monoid*. Even though we call s^* the inverse of s, we need not have $ss^* = 1$, although it is always true that $(ss^*)^2 = ss^*ss^* = ss^*$, i.e. ss^* (and s^*s for that matter) is an *idempotent*. We denote the set of all idempotents in S by

$$E(S) = \{e \in S \mid e^2 = e\}.$$

It is a nontrivial fact that if S is an inverse semigroup, then $E(S)$ is closed under multiplication and commutative. It is also clear that if $e \in E(S)$, then $e^* = e$.

Let X be a set, and let

$$\mathcal{I}(X) = \{f : U \to V \mid U, V \subset X, f \text{ bijective}\}.$$

Then $\mathcal{I}(X)$ is an inverse monoid with the operation of composition on the largest possible domain, and inverse given by function inverse; this is called the *symmetric inverse monoid on X*. Every idempotent in $\mathcal{I}(X)$ is given by Id_U for some $U \subset X$. The function Id_X is the identity for $\mathcal{I}(X)$, and the empty function is the 0 element for $\mathcal{I}(X)$. The fundamental *Wagner-Preston theorem* states that every inverse semigroup is embeddable in $\mathcal{I}(X)$ for some set X – one can think of this as analogous to the Cayley theorem for groups.

Every inverse semigroup carries a natural order structure: for $s, t \in S$ we say $s \leqslant t$ if and only if $ts^*s = s$, which is also equivalent to $ss^*t = s$. For elements $e, f \in E(S)$, we have $e \leqslant f$ if and only if $ef = e$. As usual, for $s, t \in S$, the *join* (or least upper bound) of s and t will be denoted $s \vee t$ (if it exists), and the *meet* (or greatest lower bound) of s and t will be denoted $s \wedge t$ (if it exists). For $A \subset S$, we let $A^{\uparrow} = \{t \in S \mid s \leqslant t \text{ for some } s \in A\}$ and $A^{\downarrow} = \{t \in S \mid t \leqslant s \text{ for some } s \in A\}$.

If $s, t \in S$, then we say s and t are *compatible* if $s^*t, st^* \in E(S)$, and a set $F \subset S$ is called compatible if all pairs of elements of F are compatible.

Definition 2.1. An inverse semigroup S is called *distributive* if whenever we have a compatible set $F \subset_{\text{fin}} S$, then $\bigvee_{s\in F} s$ exists in S, and for all $t \in S$ we have

$$t\left(\bigvee_{s\in F} s\right) = \bigvee_{s\in F} ts \qquad \text{and} \qquad \left(\bigvee_{s\in F} s\right)t = \bigvee_{s\in F} st.$$

In the natural partial order, the idempotents form a *meet semilattice*, which is to say that any two elements $e, f \in E(S)$ have a meet, namely ef. If $C \subset X \subset E(S)$, we say that C is a *cover* of X if for all $x \in X$ there exists $c \in C$ such that $cx \neq 0$.

In a distributive inverse semigroup each pair of idempotents has a join in addition to the meet mentioned above, but in general $E(S)$ will not have relative complements and so in general will not be a Boolean algebra. The case where $E(S)$ is a Boolean algebra is the subject of the present paper.

Definition 2.2. A *Boolean inverse monoid* is a distributive inverse monoid S with the property that $E(S)$ is a Boolean algebra, that is, for every $e \in E(S)$ there exists $e^{\perp} \in E(S)$ such that $ee^{\perp} = 0$, $e \vee e^{\perp} = 1$, and the operations $\vee, \wedge, \perp$ satisfy the laws of a Boolean algebra [GH09, Chapter 2].

Example 2.3. Perhaps the best way to think about the order structure and related concepts above is by describing them on $\mathcal{I}(X)$, which turns out to be a Boolean inverse monoid. Firstly, for $g, h \in \mathcal{I}(X)$, $g \leqslant h$ if and only if h extends g as a function. In $\mathcal{I}(X)$, two functions f and g are compatible if they agree on the intersection of their domains and their inverses agree on the intersection of their ranges. In such a situation, one can form the join

$f \vee g$ which is the union of the two functions; this will again be an element of $\mathcal{I}(X)$. Composing $h \in \mathcal{I}(X)$ with $f \vee g$ will be the same as $hf \vee hg$. Finally, $E(\mathcal{I}(X)) = \{\mathrm{Id}_U \mid U \subset X\}$ is a Boolean algebra (isomorphic to the Boolean algebra of all subsets of X) with $\mathrm{Id}_U^{\perp} = \mathrm{Id}_{U^c}$.

2.2 Étale groupoids

A *groupoid* is a small category where every arrow is invertible. If $\mathcal{G}$ is a groupoid, the set of elements $\gamma\gamma^{-1}$ is denoted $\mathcal{G}^{(0)}$ and is called the set of *units* of $\mathcal{G}$. The maps $r : \mathcal{G} \to \mathcal{G}^{(0)}$ and $d : \mathcal{G} \to \mathcal{G}^{(0)}$ defined by $r(\gamma) = \gamma\gamma^{-1}$ and $d(\gamma) = \gamma^{-1}\gamma$ are called the *range* and *source* maps, respectively.

The set $\mathcal{G}^{(2)} = \{(\gamma, \eta) \in \mathcal{G}^2 \mid r(\eta) = d(\gamma)\}$ is called the set of *composable pairs*. A *topological groupoid* is a groupoid $\mathcal{G}$ which is a topological space and for which the inverse map from $\mathcal{G}$ to $\mathcal{G}$ and the product from $\mathcal{G}^{(2)}$ to $\mathcal{G}$ are both continuous (where in the latter, the topology on $\mathcal{G}^{(2)}$ is the product topology inherited from $\mathcal{G}^2$).

We say that a topological groupoid $\mathcal{G}$ is *étale* if it is locally compact, second countable, $\mathcal{G}^{(0)}$ is Hausdorff, and the maps r and d are both local homeomorphisms. Note that an étale groupoid need not be Hausdorff. If $\mathcal{G}$ is étale, then $\mathcal{G}^{(0)}$ is open, and $\mathcal{G}$ is Hausdorff if and only if $\mathcal{G}^{(0)}$ is closed (see for example [EP16, Proposition 3.10]).

For $x \in \mathcal{G}^{(0)}$, let $\mathcal{G}(x) = \{\gamma \in \mathcal{G} \mid r(\gamma) = d(\gamma) = x\}$ – this is a group, and is called the *isotropy group at* x. A groupoid $\mathcal{G}$ is said to be *principal* if all the isotropy groups are trivial, and a topological groupoid is said to be *essentially principal* if the points with trivial isotropy groups are dense in $\mathcal{G}^{(0)}$. A topological groupoid is said to be *minimal* if for all $x \in \mathcal{G}^{(0)}$, the set $O_{\mathcal{G}}(x) = r(d^{-1}(x))$ is dense in $\mathcal{G}^{(0)}$ (the set $O_{\mathcal{G}}(x)$ is called the *orbit* of x).

If $\mathcal{G}$ is an étale groupoid, an open set $U \subset \mathcal{G}$ is called a *bisection* if $r|_U$ and $d|_U$ are both injective (and hence homeomorphisms). The set of all bisections is denoted $\mathcal{G}^{op}$ and is a distributive inverse semigroup when given the operations of setwise product and inverse. We say that an étale groupoid $\mathcal{G}$ is *ample* if the set of *compact* bisections forms a basis for the topology on $\mathcal{G}$. The set of compact bisections is called the *ample semigroup of* $\mathcal{G}$, is denoted $\mathcal{G}^a$, and is also a distributive inverse subsemigroup of $\mathcal{G}^{op}$ [LL13, Lemma 3.14]. Since $\mathcal{G}$ is second countable, $\mathcal{G}^a$ must be countable [Exe10, Corollary 4.3]. If $\mathcal{G}^{(0)}$ is compact, then the idempotent set of $\mathcal{G}^a$ is the set of all clopen sets in $\mathcal{G}^{(0)}$, and so $\mathcal{G}^a$ is a Boolean inverse monoid (see also [Ste10, Proposition 3.7] which shows that when $\mathcal{G}$ is Hausdorff and $\mathcal{G}^{(0)}$ is only locally compact, $\mathcal{G}^a$ is a Boolean inverse *semigroup*, i.e. a distributive inverse semigroup whose idempotent semilattice is a generalized Boolean algebra).

To an étale groupoid $\mathcal{G}$ one can associate C*-algebras through the theory developed by Renault [Ren80]. Let $C_c(\mathcal{G})$ denote the linear space of continuous compactly supported functions on $\mathcal{G}$. Then $C_c(\mathcal{G})$ becomes a $*$-algebra with product and involution given by

$$fg(\gamma) = \sum_{\gamma_1\gamma_2=\gamma} f(\gamma_1)g(\gamma_2), \qquad f^*(\gamma) = \overline{f(\gamma^{-1})}.$$

From this one can produce two C*-algebras $C^*(\mathcal{G})$ and $C^*_{\mathrm{red}}(\mathcal{G})$ (called the *C*-algebra of* $\mathcal{G}$ and the *reduced C*-algebra of* $\mathcal{G}$, respectively) by completing $C_c(\mathcal{G})$ in certain norms, see [Ren80, Definitions 1.12 and 2.8]. There is always a surjective $*$-homomorphism $\Lambda : C^*(\mathcal{G}) \to C^*_{\mathrm{red}}(\mathcal{G})$, and if Λ is an isomorphism we say that $\mathcal{G}$ satisfies *weak containment*. If $\mathcal{G}$ is *amenable* [ADR00], then $\mathcal{G}$ satisfies weak containment. There is an example of a case where Λ is an isomorphism for a nonamenable groupoid [Wil15], but under some conditions on $\mathcal{G}$ one has that weak containment and amenability are equivalent, see [AD16b, Theorem B].

Recall that if $B \subset A$ are both C*-algebras, then a surjective linear map $E : A \to B$ is called a *conditional expectation* if E is contractive, $E \circ E = E$, and $E(bac) = bE(a)c$ for all $b, c \in B$ and $a \in A$. Let $\mathcal{G}$ be a Hausdorff étale groupoid with compact unit space, and consider the map $E : C_c(\mathcal{G}) \to C(\mathcal{G}^{(0)})$ defined by

$$E(f) = f|_{\mathcal{G}^{(0)}}\,. \tag{1}$$

Then this map extends to a conditional expectation on both $C^*(\mathcal{G})$ and $C^*_{\mathrm{red}}(\mathcal{G})$, both denoted E. On $C^*_{\mathrm{red}}(\mathcal{G})$, E is *faithful* in the sense that if $E(a^*a) = 0$, then $a = 0$.

Let $\mathcal{G}$ be an ample étale groupoid. Both C*-algebras contain $C_c(\mathcal{G})$, and hence if U is a compact bisection, 1_U is an element of both C*-algebras. Hence we have a map $\pi : \mathcal{G}^a \to C^*(\mathcal{G})$ given by $\pi(U) = 1_U$. This map satisfies $\pi(UV) = \pi(U)\pi(V)$, $\pi(U^{-1})$, and $\pi(0) = 0$, in other words, π is a *representation* of the inverse semigroup $\mathcal{G}^a$ [Exe10].

2.3 The tight groupoid of an inverse semigroup

Let S be an inverse semigroup. A *filter* in $E(S)$ is a nonempty subset $\xi \subset E(S)$ such that

1. $0 \notin \xi$,
2. $e, f \in \xi$ implies that $ef \in \xi$, and
3. $e \in \xi, e \leqslant f$ implies $f \in \xi$.

The set of filters is denoted $\widehat{E}_0(S)$, and can be viewed as a subspace of $\{0,1\}^{E(S)}$. For $X, Y \subset_{\mathrm{fin}} E(S)$, let

$$U(X,Y) = \{\xi \in \widehat{E}_0(S) \mid X \subset \xi, Y \cap \xi = \emptyset\}.$$

sets of this form are clopen and generate the topology on $\widehat{E}_0(S)$ as X and Y vary over all the finite subsets of $E(S)$. With this topology, $\widehat{E}_0(S)$ is called the *spectrum* of $E(S)$.

A filter is called an *ultrafilter* if it is not properly contained in any other filter. The set of all ultrafilters is denoted $\widehat{E}_\infty(S)$. As a subspace of $\widehat{E}_0(S)$, $\widehat{E}_\infty(S)$ may not be closed. Let $\widehat{E}_{\text{tight}}(S)$ denote the closure of $\widehat{E}_\infty(S)$ in $\widehat{E}_0(S)$ – this is called the *tight spectrum* of $E(S)$. Of course, when $E(S)$ is a Boolean algebra, $\widehat{E}_{\text{tight}}(S) = \widehat{E}_\infty(S)$ by Stone duality [GH09, Chapter 34].
An *action* of an inverse semigroup S on a locally compact space X is a semigroup homomorphism $\alpha : S \to \mathcal{I}(X)$ such that

1. α_s is continuous for all $s \in S$,
2. the domain of α_s is open for each $s \in S$, and
3. the union of the domains of the α_s is equal to X.

If α is an action of S on X, we write $\alpha : S \curvearrowright X$. The above implies that $\alpha_{s^*} = \alpha_s^{-1}$, and so each α_s is a homeomorphism. For each $e \in E(S)$, the map α_e is the identity on some open subset D_e^α, and one easily sees that the domain of α_s is $D^\alpha_{s^*s}$ and the range of α_s is $D^\alpha_{ss^*}$, that is

$$\alpha_s : D^\alpha_{s^*s} \to D^\alpha_{ss^*}.$$

There is a natural action θ of S on $\widehat{E}_{\text{tight}}(S)$; this is referred to in [EP16] as the *standard action* of S. For $e \in E(S)$, let $D_e^\theta = \{\xi \in \widehat{E}_{\text{tight}}(S) \mid e \in \xi\} = U(\{e\}, \emptyset) \cap \widehat{E}_{\text{tight}}(S)$. For each $s \in S$ and $\xi \in D^\theta_{s^*s}$, define $\theta_s(\xi) = \{ses^* \mid e \in \xi\}^\uparrow$ – this is a well-defined homeomorphism from $D^\theta_{s^*s}$ to $D^\theta_{ss^*}$, for the details, see [Exe08].
One can associate a groupoid to an action $\alpha : S \curvearrowright X$. Let $S \times_\alpha X = \{(s, x) \in S \times X \mid x \in D^\alpha_{s^*s}\}$, and put an equivalence relation $\sim$ on this set by saying that $(s, x) \sim (t, y)$ if and only if $x = y$ and there exists some $e \in E(S)$ such that $se = te$ and $x \in D_e^\alpha$. The set of equivalence classes is denoted

$$\mathcal{G}(\alpha) = \{[s, x] \mid s \in S, x \in X\}$$

and becomes a groupoid when given the operations

$$d([s, x]) = x, \qquad r([s, x]) = \alpha_s(x),$$

$$[s, x]^{-1} = [s^*, \alpha_s(x)], \qquad [t, \alpha_s(x)][s, x] = [ts, x].$$

This is called the *groupoid of germs* of α. Note that above we are making the identification of the unit space with X, because $[e, x] = [f, x]$ for any $e, f \in E(S)$ with $x \in D_e^\alpha, D_f^\alpha$. For $s \in S$ and open set $U \subset D^\alpha_{s^*s}$ we let

$$\Theta(s, U) = \{[s, x] \mid x \in U\}$$

and endow $\mathcal{G}(\alpha)$ with the topology generated by such sets. With this topology $\mathcal{G}(\alpha)$ is an étale groupoid, sets of the above type are bisections, and if X is totally disconnected $\mathcal{G}(\alpha)$ is ample.

Let $\theta : S \curvearrowright \widehat{E}_{\text{tight}}(S)$ be the standard action, and define

$$\mathcal{G}_{\text{tight}}(S) = \mathcal{G}(\theta).$$

This is called the *tight groupoid* of S. This was defined first in [Exe08] and studied extensively in [EP16].
Let $\mathcal{G}$ be an ample étale groupoid, and consider the Boolean inverse monoid $\mathcal{G}^a$. By work of Exel [Exe10] if one uses the above procedure to produce a groupoid from $\mathcal{G}^a$, one ends up with exactly $\mathcal{G}$. In symbols,

$$\mathcal{G}_{\text{tight}}(\mathcal{G}^a) \cong \mathcal{G} \text{ for any ample étale groupoid } \mathcal{G}. \tag{2}$$

We note this result was also obtained in [Len08, Theorem 6.11] in the case where $\widehat{E}_{\text{tight}}(S) = \widehat{E}_{\infty}(S)$. In particular,

$$\mathcal{G}_{\text{tight}}(\mathcal{G}_{\text{tight}}(S)^a) \cong \mathcal{G}_{\text{tight}}(S) \text{ for all inverse semigroups } S.$$

This result can be made categorical [LL13, Theorem 3.26], and has been generalized to cases where the space of units is not even Hausdorff. This duality between Boolean inverse semigroups and ample étale groupoids falls under the broader program of noncommutative Stone duality, see [LL13] for more details.

3 C*-algebras of Boolean inverse monoids

In this section we describe the tight C*-algebra of a general inverse monoid, define the C*-algebra of a Boolean inverse monoid, and show that these two notions coincide for Boolean inverse monoids.
If S is an inverse monoid, then a *representation* of S in a unital C*-algebra A is a map $\pi : S \to A$ such that $\pi(0) = 0$, $\pi(s^*) = \pi(s)^*$, and $\pi(st) = \pi(s)\pi(t)$ for all $s, t \in S$. If π is a representation, then $C^*(\pi(E(S)))$ is a commutative C*-algebra. Let

$$\mathscr{B}_\pi = \{e \in C^*(\pi(E(S))) \mid e^2 = e = e^*\}$$

Then this set is a Boolean algebra with operations

$$e \wedge f = ef, \qquad e \vee f = e + f - ef, \qquad e^\perp = 1 - e.$$

We will be interested in a subclass of representations of S. Take $X, Y \subset_{\text{fin}} E(S)$, and define

$$E(S)^{X,Y} = \{e \in E(S) \mid e \leqslant x \text{ for all } x \in X, ey = 0 \text{ for all } y \in Y\}$$

We say that a representation $\pi : S \to A$ with A unital is *tight* if for all $X, Y, Z \subset_{\text{fin}} E(S)$ where Z is a cover of $E(S)^{X,Y}$, we have the equation

$$\bigvee_{z\in Z} \pi(z) = \prod_{x\in X} \pi(x) \prod_{y \in Y} (1 - \pi(y)). \tag{3}$$

The *tight C*-algebra* of S, denoted $C^*_{\text{tight}}(S)$, is then the universal unital C*-algebra generated by one element for each element of S subject to the relations that guarantee that the standard map from S to $C^*_{\text{tight}}(S)$ is tight. The above was all defined in [Exe08] and the interested reader is directed there for the details. It is a fact that $C^*_{\text{tight}}(S) \cong C^*(\mathcal{G}_{\text{tight}}(S))$ where the latter is the full groupoid C*-algebra (see e.g. [Exe10, Theorem 2.4]).
If S has the additional structure of being a Boolean inverse monoid, then we might wonder what extra properties π should have, in particular, what is the notion of a "join" of two partial isometries in a C*-algebra?
Let A be a C*-algebra, and suppose that S is a Boolean inverse monoid of partial isometries in A. If we have $s, t \in S$ such that $s^*t, st^* \in E(S)$, then

$$tt^*s = tt^*ss^*s = ss^*tt^*s = s(s^*t)(s^*t)^* = ss^*t$$

and if we let $a_{s,t} := s + t - ss^*t = s + t - tt^*s$, this is a partial isometry with range a_{ss^*,tt^*} and support a_{s^*s,t^*t}. A short calculation shows that $a_{s,t}$ is the least upper bound for s and t in the natural partial order, and so $a_{s,t} = s \vee t$. It is also straightforward that $r(s \vee t) = rs \vee rt$ for all $r, s, t \in S$. This leads us to the following definitions.

Definition 3.1. Let S be a Boolean inverse monoid. A *Boolean inverse monoid representation* of S in a unital C*-algebra A is a map $\pi : S \to A$ such that

1. $\pi(0) = 0$,
2. $\pi(st) = \pi(s)\pi(t)$ for all $s, t \in S$,
3. $\pi(s^*) = \pi(s)^*$ for all $s \in S$, and
4. $\pi(s \vee t) = \pi(s) + \pi(t) - \pi(ss^*t)$ for all compatible $s, t, \in S$.

Definition 3.2. Let S be a Boolean inverse monoid. Then the *universal C*-algebra of* S, denoted $C^*_B(S)$, is defined to be the universal unital C*-algebra generated by one element for each element of S subject to the relations which say that the standard map of S into $C^*_B(S)$ is a Boolean inverse monoid representation. The map π_u which takes an element s to its corresponding element in $C^*_B(S)$ will be called the *universal Boolean inverse monoid representation of* S, and we will sometimes use the notation $\delta_s := \pi_u(s)$.

The theory of tight representations was originally developed to deal with representing inverse semigroups (in which joins may not exist) inside C*-algebras, because in a C*-algebra two commuting projections always have a join. It should come as no surprise then that once we are dealing with an inverse semigroup where we <u>can</u> take joins, the representations which respect joins end up being exactly the tight representations, see [DM14, Corollary 2.3]. This is what we prove in the next proposition.

Proposition 3.3. Let S be a Boolean inverse monoid. Then a map $\pi : S \to A$ is a Boolean inverse monoid representation of S if and only if π is a tight representation.

Proof. Suppose that π is a Boolean inverse monoid representation of S. Then when restricted to $E(S)$, π is a Boolean algebra homomorphism into $\mathscr{B}_\pi$, and so by [Exe08, Proposition 11.9], π is a tight representation.
On the other hand, suppose that π is a tight representation, and first suppose that $e, f \in E(S)$. Then the set $\{e, f\}$ is a cover for $E(S)^{\{e\vee f\},\emptyset}$, so

$$\pi(e) \vee \pi(f) = \pi(e \vee f).$$

Now let $s, t \in S$ be compatible, so that $s^*t = t^*s$ and $st^* = ts^*$ are both idempotents, and we have

$$s^*st^*t = s^*ts^*t = s^*t.$$

Since $(s \vee t)^*(s \vee t) = s^*s \vee t^*t$, we have

$$\begin{aligned}
\pi(s \vee t) &= \pi(s \vee t)\pi(s^*s \vee t^*t) \\
&= \pi(s \vee t)(\pi(s^*s) + \pi(t * t) - \pi(s^*st^*t) \\
&= \pi(ss^*s \vee ts^*s) + \pi(st^*t \vee tt^*t) - \pi(ss^*st^*t \vee tt^*ts^*s)) \\
&= \pi(s \vee st^*s) + \pi(ts^*t \vee t) - \pi(st^*t \vee ts^*s) \\
&= \pi(s) + \pi(t) - \pi(ss^*t)
\end{aligned}$$

where the last line follows from the facts that $st^*s \leqslant s$, $ts^*t \leqslant t$ and $ts^*s = st^*t = ss^*t = tt^*s$. ☐

We have the following consequence of the proof of the above proposition.

COROLLARY 3.4. Let S be a Boolean inverse monoid. Then a map $\pi : S \to A$ is a Boolean inverse monoid representation of S if and only if it is a representation and for all $e, f \in E(S)$ we have $\pi(e \vee f) = \pi(e) + \pi(f) - \pi(ef)$.

We now have the following.

THEOREM 3.5. Let S be a Boolean inverse monoid. Then

$$C^*_B(S) \cong C^*_{\text{tight}}(S) \cong C^*(\mathcal{G}_{\text{tight}}(S)).$$

In what follows, we will be studying traces on C*-algebras arising from Boolean inverse monoids. However, many of our examples will actually arise from inverse monoids which are not distributive, and so the Boolean inverse monoid in question will actually be $\mathcal{G}_{\text{tight}}(S)^a$, see (2). The map from S to $\mathcal{G}_{\text{tight}}(S)^a$ defined by

$$s \mapsto \Theta(s, D^\theta_{s^*s})$$

may fail to be injective, and so we cannot say that a given inverse monoid can be embedded in a Boolean inverse monoid. The obstruction arises from the following situation: suppose S is an inverse semigroup and that we have $e, f \in E(S)$ such that $e \leqslant f$ and for all $0 \neq k \leqslant f$ we have $ek \neq 0$, in other words, $\{e\}$ is a cover for $\{f\}^\downarrow$. In such a situation, we say that e is *dense* in

f^2, and by (3) we must have that $\pi(e) = \pi(f)$ (see also [Exe09] and [Exe08, Proposition 11.11]). For most of our examples, we will be considering inverse semigroups which have faithful tight representations, though we consider one which does not.

We close this section by recording some consequences of Theorem 3.5. The tight groupoid and tight C*-algebra of an inverse semigroup were extensively studied in [EP16] and [Ste16], where they gave conditions on S which imply that $C^*_{\text{tight}}(S)$ is simple and purely infinite. We first recall some definitions from [EP16].

Definition 3.6. Let S be an inverse semigroup, let $s \in S$ and $e \leqslant s^*s$. Then we say that

1. e is *fixed* by s if $se = e$, and

2. e is *weakly fixed* by s if for all $0 \neq f \leqslant e$, $fsfs^* \neq 0$.

Denote by $\mathcal{J}_s := \{e \in E(S) \mid se = e\}$ the set of all fixed idempotents for $s \in S$. We note that an inverse semigroup for which $\mathcal{J}_s = \{0\}$ for all $s \notin E(S)$ is called *E^*-unitary.*

Theorem 3.7. Let S be an inverse semigroup. Then

1. $\mathcal{G}_{\text{tight}}(S)$ is Hausdorff if and only if $\mathcal{J}_s$ has a finite cover for all $s \in S$. [EP16, Theorem 3.16]

2. If $\mathcal{G}_{\text{tight}}(S)$ is Hausdorff, then $\mathcal{G}_{\text{tight}}(S)$ is essentially principal if and only if for every $s \in S$ and every $e \in E(S)$ weakly fixed by s, there exists a finite cover for $\{e\}$ by fixed idempotents. [EP16, Theorem 4.10]

3. $\mathcal{G}_{\text{tight}}(S)$ is minimal if and only if for every nonzero $e, f \in E(S)$, there exist $F \subset_{\text{fin}} S$ such that $\{esfs^* \mid s \in F\}$ is a cover for $\{e\}$.[EP16, Theorem 5.5]

We translate the above to the case where S is a Boolean inverse monoid.

Proposition 3.8. Let S be a Boolean inverse monoid. Then

1. $\mathcal{G}_{\text{tight}}(S)$ is Hausdorff if and only if for all $s \in S$, there exists an idempotent e_s with $se_s = e_s$ such that if e is fixed by s, then $e \leqslant e_s$.

2. If $\mathcal{G}_{\text{tight}}(S)$ is Hausdorff, then $\mathcal{G}_{\text{tight}}(S)$ is essentially principal if and only if for every $s \in S$, e weakly fixed by s implies e is fixed by s.

3. $\mathcal{G}_{\text{tight}}(S)$ is minimal if and only if for every nonzero $e, f \in E(S)$, there exist $F \subset_{\text{fin}} S$ such that $e \leqslant \bigvee_{s \in F} sfs^*$.

---

[2] This is the terminology used in [Exe08, Definition 11.10] and [Exe09], though in [LS14, Section 6.3] such an e is called *essential* in f.

Proof. Statements 2 and 3 are easy consequences of taking the joins of the finite covers mentioned. Statement 1 is central to what follows, and is proven in Lemma 4.2. □

If an étale groupoid $\mathcal{G}$ is Hausdorff, then $C^*(\mathcal{G})$ is simple if and only if $\mathcal{G}$ is essentially principal, minimal, and satisfies weak containment, see [BCFS14] (also see [ES15] for a discussion of amenability of groupoids associated to inverse semigroups).

4 INVARIANT MEANS AND TRACES

In this section we consider invariant means on Boolean inverse monoids, and show that such functions always give rise to traces on the associated C*-algebras. This definition is from [KLLR16].

DEFINITION 4.1. Let S be a Boolean inverse monoid. A nonzero function $\mu : E(S) \to [0, \infty)$ will be called an *invariant mean* if

1. $\mu(s^*s) = \mu(ss^*)$ for all $s \in S$
2. $\mu(e \vee f) = \mu(e) + \mu(f)$ for all $e, f \in E(S)$ such that $ef = 0$.

If in addition $\mu(1) = 1$, we call μ a *normalized invariant mean.* An invariant mean μ will be called *faithful* if $\mu(e) = 0$ implies $e = 0$. We will denote by $M(S)$ the affine space of all normalized invariant means on S.

We make an important assumption on the Boolean inverse monoids we consider here. This assumption is equivalent to the groupoid $\mathcal{G}_{\text{tight}}(S)$ being Hausdorff [EP16, Theorem 3.16].[3]

$$\text{For every } s \in S, \text{ the set } \mathcal{J}_s = \{e \in E(S) \mid se = e\} \text{ admits a finite cover.} \tag{H}$$

The next lemma records straightforward consequences of condition (H) when S happens to be a Boolean inverse monoid.

LEMMA 4.2. Let S be Boolean inverse monoid which satisfies condition (H). Then,

1. for each $s \in S$ there is an idempotent e_s such that for any finite cover C of $\mathcal{J}_s$,
$$e_s = \bigvee_{c \in C} c. \tag{4}$$
and $\mathcal{J}_s = \mathcal{J}_{e_s}$,

[3] In [Sta15], we define condition (H) for another class of semigroups, namely the right LCM semigroups. Right LCM semigroups and inverse semigroups are related, but the intersection of their classes is empty (because right LCM semigroups are left cancellative and we assume that our inverse semigroups have a zero element). We note that a right LCM semigroup P satisfies condition (H) in the sense of [Sta15] if and only if its left inverse hull $I_l(P)$ satisfies condition (H) in the sense of the above.

2. $e_{s^*} = e_s$ for all $s \in S$,

3. $e_{st} \leqslant ss^*, t^*t$ for all $s, t \in S$, and

4. $e_{s^*t}e_{t^*r} \leqslant e_{s^*r}$ for all $s, t, r \in S$.

Proof. To show the first statement, we need to show that any two covers give the same join. If $\mathcal{J}_s = \{0\}$, there is nothing to do. So suppose that $0 \neq e \in \mathcal{J}_s$, suppose that C is a cover for $\mathcal{J}_s$, and let $e_C = \bigvee_{c \in C} c$. Indeed, the element $ee_C^{\perp}$ must be in $\mathcal{J}_s$, and since it is orthogonal to all elements of C and C is a cover, $ee_C^{\perp}$ must be 0. Hence we have

$$e = ee_C \vee ee_C^{\perp} = ee_C$$

and so $e \leqslant e_C$. Now if K is another cover for $\mathcal{J}_s$ with join e_K and $k \in K$, we must have that $k \leqslant e_C$, and so $e_K \leqslant e_C$. Since the argument is symmetric, we have proven the first statement.
To prove the second statement, if $e \in \mathcal{J}_s$ then we have

$$ses^* = es^* = (se)^* = e$$

and so

$$s^*e = s^*(ses^*) = es^*ss^* = es^* = (se)^* = e$$

and again by symmetry we have $\mathcal{J}_s = \mathcal{J}_{s^*}$ and so $e_s = e_{s^*}$.
To prove the third statement, we notice

$$ss^*e_{st} = ss^*ste_{st} = ste_{st} = e_{st}$$

$$e_{st}t^*t = ste_{st}t^*t = stt^*te_{st} = ste_{st} = e_{st}.$$

For the fourth statement, we calculate (using 2)

$$\begin{aligned} e_{s^*t}e_{t^*r} &= s^*te_{s^*t}e_{t^*r} = s^*tt^*re_{s^*t}e_{t^*r} \\ &= s^*tt^*rr^*te_{s^*t}e_{t^*r} = s^*rr^*te_{s^*t}e_{t^*r} \\ &= s^*re_{s^*t}e_{t^*r} \end{aligned}$$

hence $e_{s^*t}e_{t^*r} \leqslant s^*r$ and so $e_{s^*t}e_{t^*r} \leqslant e_{s^*r}$. $\square$

In what will be a crucial step to obtaining a trace from an invariant mean, we now obtain a relationship between e_{st} and e_{ts}.

Lemma 4.3. Let S be Boolean inverse monoid which satisfies condition (H). Then for all $s, t \in S$, we have that $s^*e_{st}s = e_{ts}$.

Proof. Suppose that $e \in \mathcal{J}_{ts}$. Then $tse = e$, and so

$$(st)ses^* = ses^*$$

hence $ses^* \in \mathcal{J}_{st}$. If C is a cover of $\mathcal{J}_{st}$ and $f \in \mathcal{J}_{ts}$, there must exist $c \in C$ such that $c(sfs^*) \neq 0$. Hence

$$\begin{aligned} css^*sfs^* &\neq 0 \\ ss^*csfs^* &\neq 0 \\ s^*csf &\neq 0 \end{aligned}$$

and so we see that s^*Cs is a cover for $\mathcal{J}_{ts}$. By Lemma 4.2,

$$e_{ts} = \bigvee_{c \in C} s^*cs = s^* \left(\bigvee_{c \in C} c \right) s = s^*e_{st}s.$$

□

Lemma 4.3 and Lemma 4.2.3 imply that for all $s, t \in S$ and all $\mu \in M(S)$, we have $\mu(e_{st}) = \mu(e_{ts})$.

REMARK 4.4. We are thankful to Ganna Kudryavtseva for pointing out to us that the proofs Lemmas 4.2 and 4.3 can be simplified by using the fact from [KL14, Theorem 8.20] that a Boolean inverse monoid S satisfies condition (H) if and only if every pair of elements in S has a meet (see also [Ste10, Proposition 3.7] for another wording of this fact). From this, one can see that for all $s \in S$ we have

$$e_s = s \wedge (s^*s) = s \wedge (ss^*).$$

DEFINITION 4.5. Let A be a C*-algebra. A bounded linear functional $\tau : A \to \mathbb{C}$ is called a *trace* if

1. $\tau(a^*a) \geq 0$ for all $a \in A$,
2. $\tau(ab) = \tau(ba)$ for all $a, b \in A$.

A trace τ is said to be *faithful* if $\tau(a^*a) > 0$ for all $a \neq 0$. A trace τ on a unital C*-algebra is called a *tracial state* if $\tau(1) = 1$. The set of all tracial states of a C*-algebra A is denoted $T(A)$.

We are now able to define a trace on $C^*_B(S)$ for each $\mu \in M(S)$.

PROPOSITION 4.6. Let S be Boolean inverse monoid which satisfies condition (H), and let $\mu \in M(S)$. Then there is a trace τ_μ on $C^*_B(S)$ such that

$$\tau_\mu(\delta_s) = \mu(e_s) \qquad \text{for all } s \in S.$$

If μ is faithful, then the restriction of τ_μ to $C^*_{\text{red}}(\mathcal{G}_{\text{tight}}(S))$ is a faithful trace.

Proof. We define τ_μ to be as above on the generators δ_s of $C^*_B(S)$, and extend it to B :=span$\{\delta_s \mid s \in S\}$, a dense $*$-subalgebra of $C^*_B(S)$.

We first show that $\tau_\mu(\delta_s\delta_t) = \tau_\mu(\delta_t\delta_s)$. Indeed, by Lemmas 4.2 and 4.3, we have

$$\begin{aligned}\tau_\mu(\delta_s\delta_t) &= \mu(e_{st}) = \mu(e_{st}ss^*) = \mu(e_{st}ss^*e_{st}) = \mu((e_{st}s)(e_{st}s)^*)\\ &= \mu((e_{st}s)^*(e_{st}s)) = \mu(s^*e_{st}s) = \mu(e_{ts}) = \tau_\mu(\delta_t\delta_s).\end{aligned}$$

Since τ_μ is extended linearly to B, we have that $\tau_\mu(ab) = \tau_\mu(ba)$ for all $a, b \in B$. Let F be a finite index set and take $x = \sum_{i\in F} a_i\delta_{s_i}$ in B. We will show that $\tau_\mu(x^*x) \geq 0$. For $i, j \in F$, we let $e_{ij} = e_{s_i^*s_j}$ and note that $e_{ij} = e_{ji}$. We calculate:

$$\begin{aligned}x^*x &= \left(\sum_{s\in S}\overline{a_i}\delta_{s_i^*}\right)\left(\sum_{j\in F}a_j\delta_{s_j}\right)\\ &= \sum_{i,j\in F}\overline{a_i}a_j\delta_{s_i^*s_j}\\ \tau_\mu(x^*x) &= \sum_{i,j\in F}\overline{a_i}a_j\mu(e_{ij})\\ &= \sum_{i\in F}|a_i|^2\mu(e_{ii}) + \sum_{i,j\in F, i\neq j}(\overline{a_i}a_j + \overline{a_j}a_i)\mu(e_{ij}).\end{aligned}$$

We will show that this sum is positive by using an orthogonal decomposition of the e_{ij}. Let $F^2_{\neq} = \{\{i,j\} \subset F \mid i \neq j\}$, and let $D(F^2_{\neq}) = \{(A,B) \mid A\cup B = F^2_{\neq}, A\cap B = \emptyset\}$. For $a = \{i,j\} \in F^2_{\neq}$, let $e_a = e_{ij}$. We have

$$e_{ij} = e_{ij}\bigvee_{(A,B)\in D(F^2_{\neq})}\left(\prod_{a\in A, b\in B}e_ae_b^\perp\right)$$

where the join is an orthogonal join. Of course, the above is only nonzero when $\{i,j\} \in A$. We also notice that

$$e_{ii} \geqslant \bigvee_{\substack{(A,B)\in D(F^2_{\neq})\\ i\in\cup A}}\left(\prod_{a\in A, b\in B}e_ae_b^\perp\right)$$

and so $\tau_\mu(x^*x)$ is larger than a linear combination of terms of the form $\mu\left(\prod_{a\in A,b\in B}e_ae_b^\perp\right)$ for partitions (A,B) of $F^2_{\neq}$: specifically, $\tau_\mu(x^*x)$ is greater than or equal to

$$\sum_{(A,B)\in D(F^2_{\neq})}\left[\left(\sum_{i\in\cup A}|a_i|^2 + \sum_{a=\{j,k\}\in A}(\overline{a_i}a_j + \overline{a_j}a_i)\right)\mu\left(\prod_{a\in A,b\in B}e_ae_b^\perp\right)\right] \quad (5)$$

If a term $\prod_{a\in A,b\in B} e_a e_b^\perp$ is not zero, then we claim that the relation

$$i \sim j \text{ if and only if } i = j \text{ or } \{i,j\} \in A$$

is an equivalence relation on $\cup A$. Indeed, suppose that $i,j,k \in \cup A$ are all pairwise nonequal and $\{i,j\},\{j,k\} \in A$. By Lemma 4.2.4, $e_{ij}e_{jk} \leqslant e_{ik}$ and since the product is nonzero, we must have that $\{i,k\} \in A$. Writing $[\cup A]$ for the set of equivalence classes, we have

$$\begin{aligned}\sum_{i\in\cup A}|a_i|^2 + \sum_{a=\{j,k\}\in A}(\overline{a_i}a_j + \overline{a_j}a_i) &= \sum_{C\in[\cup A]}\left(\sum_{i\in C}|a_i|^2 + \sum_{\substack{i,j\in C\\ i\neq j}}(\overline{a_i}a_j + \overline{a_j}a_i)\right)\\ &= \sum_{C\in[\cup A]}\left|\sum_{i\in C}a_i\right|^2.\end{aligned}$$

Hence, $\tau_\mu(x^*x) \geq 0$, and τ_μ is positive on B. Hence, τ_μ extends to a trace on $C^*_B(S)$.

The above calculation shows that if μ is faithful, then τ_μ is faithful on B. A short calculation shows that $E(\delta_s) = \delta_{e_s}$, where E is as in (1). Furthermore, it is clear that on B we have that $\tau_\mu = \tau_\mu \circ E$, and so we will show that τ_μ is faithful on $C^*_{\mathrm{red}}(\mathcal{G}_{\mathrm{tight}}(S))$ if we show that $\tau_\mu(a) > 0$ for all nonzero positive $a \in C(\widehat{E}_{\mathrm{tight}}(S))$. If $a \in C(\widehat{E}_{\mathrm{tight}}(S))$ is positive, then it is bounded above zero on some clopen set given by D_e for some $e \in E(S)$. Hence, $\tau_\mu(a) \geq \tau_\mu(\delta_e) = \mu(e)$ which must be strictly positive because μ is faithful. ☐

We now show that given a trace on $C^*_B(S)$ we can construct an invariant mean on S.

PROPOSITION 4.7. Let S be Boolean inverse monoid, let $\pi_u : S \to C^*_B(S)$ be the universal Boolean monoid representation of S, and take $\tau \in T(C^*_B(S))$. Then the map $\mu_\tau : E(S) \to [0,\infty)$ defined by

$$\mu_\tau(e) = \tau(\pi_u(e)) = \tau(\delta_e)$$

is a normalized invariant mean on S. If τ is faithful then so is μ_τ.

Proof. That μ_τ takes positive values follows from τ being positive. We have

$$\begin{aligned}\mu_\tau(s^*s) &= \tau(\pi_u(s^*s)) = \tau(\pi_u(s^*)\pi_u(s))\\ &= \tau(\pi_u(s)\pi_u(s^*)) = \tau(\pi_u(ss^*))\\ &= \mu_\tau(ss^*).\end{aligned}$$

Also, if $e,f \in E(S)$ with $ef = 0$, then

$$\begin{aligned}\mu_\tau(e\vee f) &= \tau(\pi_u(e\vee f)) = \tau(\pi_u(e) + \pi_u(f))\\ &= \tau(\pi_u(e)) + \tau(\pi_u(f))\\ &= \mu_\tau(e) + \mu_\tau(f).\end{aligned}$$

If τ is faithful and $e \neq 0$, $\tau(\delta_e) > 0$ because δ_e is positive and nonzero, and so μ_τ is faithful. $\square$

Proposition 4.8. Let S be Boolean inverse monoid which satisfies condition (H). Then the map

$$\mu \mapsto \tau_\mu \mapsto \mu_{\tau_\mu}$$

is the identity on $M(S)$.

Proof. This is immediate, since if $\mu \in M(S)$ and $e \in E(S)$ we have

$$\mu_{\tau_\mu}(e) = \tau_\mu(\pi_u(e)) = \tau_\mu(\delta_e) = \mu(e).$$

$\square$

Given the above, one might wonder under which circumstances we have that $T(C^*_B(S)) \cong M(S)$. This is not true in the general situation – take for example S to be the group $\mathbb{Z}_2 = \{1, -1\}$ with a zero element adjoined – this is a Boolean inverse monoid. Here $M(S)$ consists of one element, namely the function which takes the value 1 on 1 and the value 0 on the zero element. The C*-algebra of S is the group C*-algebra of $\mathbb{Z}_2$, which is isomorphic to $\mathbb{C}^2$, a C*-algebra with many traces (taking the dot product of an element of $\mathbb{C}^2$ with any nonnegative vector whose entries add to 1 determines a normalized trace on $\mathbb{C}^2$).
One can still obtain this isomorphism using the following.

Definition 4.9. Let $\mathcal{G}$ be an étale groupoid. A regular Borel probability measure ν on $\mathcal{G}^{(0)}$ is called *$\mathcal{G}$-invariant* if for every bisection U one has that $\nu(r(U)) = \nu(d(U))$. The affine space of all regular $\mathcal{G}$-invariant Borel probability measures is denoted $IM(\mathcal{G})$.

The following is a special case of [KR06, Proposition 3.2].

Theorem 4.10. (cf [KR06, Proposition 3.2]) Let $\mathcal{G}$ be a Hausdorff principal étale groupoid with compact unit space. Then

$$T(C^*_{\mathrm{red}}(\mathcal{G})) \cong IM(\mathcal{G})$$

For $\tau \in T(C^*_{\mathrm{red}}(\mathcal{G}))$ the image of τ under the above isomorphism is the regular Borel probability measure ν whose existence is guaranteed by the Riesz representation theorem applied to the positive linear functional on $C(\mathcal{G}^{(0)})$ given by restricting τ.

For a proof of Theorem 4.10 in the above form, see [Put, Theorem 3.4.5].
For us, the groupoid $\mathcal{G}_{\mathrm{tight}}(S)$ satisfies all of the conditions in Theorem 4.10, except possibly for being principal. Also note that in the general case, $C^*_{\mathrm{red}}(\mathcal{G}_{\mathrm{tight}})$ may not be isomorphic to $C^*_B(S)$. So if we restrict our attention to Boolean inverse monoids which have principal tight groupoids and for which $C^*_{\mathrm{red}}(\mathcal{G}_{\mathrm{tight}}(S)) \cong C^*_B(S)$ (that is to say, Boolean inverse monoids for which $\mathcal{G}_{\mathrm{tight}}(S)$ satisfies weak containment), we can obtain the desired isomorphism. While this may seem like a restrictive set of assumptions, they are all satisfied for the examples we consider here.

Proposition 4.11. Let S be Boolean inverse monoid which satisfies condition (H), and suppose $\nu \in IM(\mathcal{G}_{\text{tight}}(S))$. Then the map $\eta_\nu : E(S) \to [0, \infty)$ defined by

$$\eta_\nu(e) = \nu(D_e^\theta)$$

is a normalized invariant mean on S. The map that sends $\nu \mapsto \eta_\nu$ is an affine isomorphism of $IM(\mathcal{G}_{\text{tight}}(S))$ and $M(S)$.

Proof. That $\eta_\nu(s^*s) = \eta_\nu(ss^*)$ follows from invariance of ν applied to the bisection $\Theta(s, D_{s^*s})$, and that η_ν is additive over orthogonal joins follows from the fact that ν is a measure. This map is clearly affine. Suppose that $\eta_\nu = \eta_\kappa$ for $\nu, \kappa \in IM(\mathcal{G}_{\text{tight}}(S))$. Then ν, κ agree on all sets of the form D_e^θ, and since these sets generate the topology on $\widehat{E}_{\text{tight}}(S)$, ν and κ agree on all open sets. Since they are regular Borel probability measures they must be equal, and so $\nu \mapsto \eta_\nu$ is injective.

To get surjectivity, let μ be an invariant mean, and let τ_μ be as in Proposition 4.6. Then restricting τ_μ to $C(\widehat{E}_{\text{tight}}(S))$ and invoking the Riesz representation theorem gives us a regular invariant probability measure ν on $\widehat{E}_{\text{tight}}(S)$, and we must have $\eta_\nu = \mu$. □

Corollary 4.12. Let $\mathcal{G}$ be an ample Hausdorff groupoid. Then $IM(\mathcal{G}) \cong M(\mathcal{G}^a)$.

So the invariant means on the ample semigroup of an ample Hausdorff groupoid are in one-to-one correspondence with the $\mathcal{G}$-invariant measures.

Theorem 4.13. Let S be Boolean inverse monoid which satisfies condition (H). Suppose that $\mathcal{G}_{\text{tight}}(S)$ is principal, and that $C^*_{\text{red}}(\mathcal{G}_{\text{tight}}(S)) \cong C^*_B(S)$. Then

$$T(C^*_B(S)) \cong M(S)$$

via the map which sends τ to μ_τ as in Proposition 4.7. In addition, both are isomorphic to $IM(\mathcal{G}_{\text{tight}}(S))$.

Proof. This follows from Theorem 4.10 and Proposition 4.11. □

There are many results in the literature concerning traces which now apply to our situation.

Corollary 4.14. Let S be Boolean inverse monoid which satisfies condition (H). If S admits a faithful invariant mean, then $C^*_{\text{red}}(\mathcal{G}_{\text{tight}}(S))$ is stably finite. If in addition $\mathcal{G}_{\text{tight}}(S)$ satisfies weak containment, $C^*_B(S)$ is stably finite.

Proof. If μ is a faithful invariant mean, then after normalizing one obtains a faithful trace on $C^*_{\text{red}}(\mathcal{G}_{\text{tight}}(S))$ by Proposition 4.6. Now the result is standard, see for example [LLR00, Exercise 5.2]. □

Corollary 4.15. Let S be Boolean inverse monoid which satisfies condition (H). If $C^*_B(S)$ is stably finite and exact, then S has an invariant mean.

Proof. This is a consequence of the celebrated result of Haagerup [Haa91] when applied to Proposition 4.7 □

For the undefined terms above, we direct the interested reader to [BO08]. We also note that exactness of $C^*_B(S)$ has recently been considered in [Li16] and [AD16a].

5 Examples

5.1 AF inverse monoids

This is a class of Boolean inverse monoids introduced in [LS14] motivated by the construction of AF C*-algebras from Bratteli diagrams.
A *Bratteli diagram* is an infinite directed graph $B = (V, E, r, s)$ such that

1. V can be written as a disjoint union of finite sets $V = \cup_{n\geq 0} V_n$
2. V_0 consists of one element v_0, called the *root*,
3. for all edges $e \in E$, $s(e) \in V_i$ implies that $r(e) \in V_{i+1}$ for all $i \geq 0$, and
4. for all $i \geq 1$ and all $v \in V_i$, both $r^{-1}(v)$ and $s^{-1}(v)$ are finite and nonempty.

We also denote $s^{-1}(V_i) := E_i$, so that $E = \cup_{n\geq 0} E_n$. Let E^* be the set of all finite paths in B, including the vertices (treated as paths of length zero). For $v, w \in V \cup E$, let vE^* denote all the paths starting with v, let E^*w be all the paths ending with w, and let vE^*w be all the paths starting with v and ending with w.
Given a Bratteli diagram $B = (V, E, r, s)$ we construct a C*-algebra as follows. We let

$$A_0 = \mathbb{C}$$

$$A_1 = \bigoplus_{v\in V_1} \mathbb{M}_{|r^{-1}(v)|},$$

and define $k_1(v) = |r^{-1}(v)|$ for all $v \in V_1$. For an integer $i > 1$ and $v \in V_i$, let

$$k_i(v) = \sum_{\gamma\in r^{-1}(v)} k_{i-1}(s(\gamma)). \tag{6}$$

Define

$$A_i = \bigoplus_{v\in V_i} \mathbb{M}_{k_i(v)}$$

Now for all $i \geq 0$, one can embed $A_i \hookrightarrow A_{i+1}$ by viewing, for each $v \in V_{i+1}$

$$\bigoplus_{\gamma\in r^{-1}(v)} \mathbb{M}_{k_i(s(\gamma))} \subset \mathbb{M}_{k_{i+1}(v)}$$

where the algebras in the direct sum are orthogonal summands along the diagonal in $\mathbb{M}_{k_{i+1}(v)}$. So $A_0 \hookrightarrow A_1 \hookrightarrow A_1 \hookrightarrow \cdots$ can be viewed as an increasing union of finite dimensional C*-algebra, all of which can be realized as subalgebras of $\mathcal{B}(\mathcal{H})$ for the same $\mathcal{H}$, and so we can form the norm closure of the union

$$A_B := \overline{\bigcup_{n\geq 0} A_n}.$$

This C*-algebra is what is known as an *AF algebra*, and every unital AF algebra arises this way from some Bratteli diagram.
The AF algebra A_B can always be described as the C*-algebra of a principal groupoid derived from B, see [Ren80] and [ER06]. We reproduce this construction here. Let X_B denote the set of all infinite paths in B which start at the root. When given the product topology from the discrete topologies on the E_n, this is a compact Hausdorff totally disconnected space. For $\alpha \in v_0E^*$, we let $C(\alpha) = \{x \in X_B \mid x_i = \alpha_i \text{ for all } i = 0, \ldots, |\alpha| - 1\}$. Sets of this form are clopen and form a basis for the topology on X_B. For $n \in \mathbb{N}$, let

$$\mathcal{R}_B^{(n)} = \{(x,y) \in X \times X \mid x_i = y_i \text{ for all } i \geq n+1\}$$

so a pair of infinite paths (x, y) is in $\mathcal{R}_B^{(n)}$ if and only if x and y agree after the vertices on level n. Clearly, $\mathcal{R}_B^{(n)} \subset \mathcal{R}_B^{(n+1)}$, and so we can form their union

$$\mathcal{R}_B = \bigcup_{n\in\mathbb{N}} \mathcal{R}_B^{(n)}.$$

This is an equivalence relation, known as *tail equivalence* on X_B. For $v \in V \setminus \{v_0\}$ and $\alpha, \beta \in v_0E^*v$, define

$$C(\alpha, \beta) = \{(x,y) \in \mathcal{R}_B \mid x \in C(\alpha), y \in C(\beta)\}$$

sets of this type form a basis for a topology on $\mathcal{R}_B$, and with this topology $\mathcal{R}_B$ is a principal Hausdorff étale groupoid with unit space identified with X_B, and

$$C^*(\mathcal{R}_B) \cong C^*_{\text{red}}(\mathcal{R}_B) \cong A_B.$$

In [LS14], a Boolean inverse monoid is constructed from a Bratteli diagram, mirroring the above construction. We will present this Boolean inverse monoid in a slightly different way which may be enlightening. Let $B = (V, E, r, s)$ be a Bratteli diagram. Let S_0 be the Boolean inverse monoid (in fact, Boolean algebra) $\{0, 1\}$. For each $i \geq 1$, let

$$S_i = \bigoplus_{v\in V_i} \mathcal{I}(v_0E^*v)$$

where as in Section 2.1, $\mathcal{I}(X)$ denotes the set of partially defined bijections on X.

If $v \in V_{i+1}$ and $\gamma \in r^{-1}(v)$ then one can view $\mathcal{I}(v_0E^*\gamma)$ as a subset of $\mathcal{I}(v_0E^*v)$, and if $\eta \in r^{-1}(v)$ with $\gamma \neq \eta$, $\mathcal{I}(v_0E^*\gamma)$ and $\mathcal{I}(v_0E^*\eta)$ are orthogonal. Furthermore, $\mathcal{I}(v_0E^*\gamma)$ can be identified with $\mathcal{I}(v_0E^*s(\gamma))$ Hence the direct sum over $r^{-1}(v)$ can be embedded into $\mathcal{I}(v_0E^*v)$:

$$\bigoplus_{\gamma\in r^{-1}(v)} \mathcal{I}(v_0E^*s(\gamma)) \hookrightarrow \mathcal{I}(v_0E^*v). \tag{7}$$

This allows us to embed $S_i \hookrightarrow S_{i+1}$

$$\bigoplus_{v\in V_i} \mathcal{I}(v_0E^*v) \hookrightarrow \bigoplus_{w\in V_{i+1}} \mathcal{I}(v_0E^*w)$$

where an element ϕ in a summand $\mathcal{I}(v_0E^*v)$ gets sent to $|s^{-1}(v)|$ summands on the right, one for each $\gamma \in s^{-1}(v)$: ϕ will be sent to the summand inside $\mathcal{I}(v_0E^*s(\gamma))$ corresponding to v in left hand side of the embedding from (7). We then define

$$I(B) = \varinjlim(S_i \hookrightarrow S_{i+1})$$

This is a Boolean inverse monoid [LS14, Lemma 3.13]. As a set $I(B)$ is the union of all the S_i, viewed as an increasing union via the identifications above. In [LS14, Remark 6.5], it is stated that the groupoid one obtains from $I(B)$ (i.e., $\mathcal{G}_{\text{tight}}(I(B))$) is exactly tail equivalence. We provide the details of that informal discussion here.
We will describe the ultrafilters in $E(I(B))$, a Boolean algebra. For $v \in V_i$ and a path $\alpha \in v_0E^*v$, let $e_\alpha = \mathrm{Id}_{\{\alpha\}} \in \mathcal{I}(v_0E^*v)$. As v ranges over all of V_i and α ranges over all of v_0E^*v, these idempotents form a orthogonal decomposition of the identity of $I(B)$. Hence, given an ultrafilter ξ and $i > 0$ there exists one and only one path, say $\alpha_\xi^{(i)}$ ending at level i with $e_{\alpha_\xi^{(i)}} \in \xi$. Furthermore, if $j > i$, we must have that $\alpha_\xi^{(i)}$ is a prefix of $\alpha_\xi^{(j)}$, because products in an ultrafilter cannot be zero. So for $x \in X_B$, if we define

$$\xi_x = \{e_\alpha \mid \alpha \text{ is a prefix of } x\}$$

then we have that

$$\widehat{E}_\infty(I(B)) = \{\xi_x \mid x \in X_B\}$$

By [EP16, Proposition 2.6], the set

$$\{U(\{e_\alpha\}, \emptyset) \mid \alpha \text{ is a prefix of } x\}$$

is a neighbourhood basis for ξ_x. The map $\lambda : X_B \to \widehat{E}_\infty(I(B))$ given by $\lambda(x) = \xi_x$ is a bijection, and since $U(\{e_\alpha\}, \emptyset) = \lambda(C(\alpha))$, it is a homeomorphism. If $\phi \in S_i$ such that $\phi^*\phi \in \xi_x$, then we must have that one component of ϕ is in $\mathcal{I}(v_0E^*r(x_i))$, and we must have that

$$\theta_\phi(\xi_x) = \xi_{\phi(x_0x_1\ldots x_i)x_{i+1}x_{i+2}\ldots} \tag{8}$$

Finally, we claim that $\mathcal{R}_B$ is isomorphic to $\mathcal{G}_{\text{tight}}(I(B))$. We define a map

$$\Phi : \mathcal{G}_{\text{tight}}(I(B)) \to \mathcal{R}_B$$

$$\Phi([\phi, \xi_x]) \mapsto (\phi(x_0 x_1 \dots x_i) x_{i+1} x_{i+2} \dots, x)$$

where ϕ and x are as in (8). If $\Phi([\phi, \xi_x]) = \Phi([\psi, \xi_y])$, then clearly we must have $\xi_x = \xi_y$. We must also have that $\phi, \psi \in S_i$, and $\phi e_{x_0 x_1 \dots x_i} = \psi e_{x_0 x_1 \dots x_i}$, hence $[\phi, \xi_x] = [\psi, \xi_y]$. It is straightforward to verify that Φ is surjective and bicontinuous, and so $\mathcal{R}_B \cong \mathcal{G}_{\text{tight}}(I(B))$. Since they are both étale, their C*-algebras must be isomorphic. Hence with the above discussion, we have proven the following.

Theorem 5.1. Let B be a Bratteli diagram. Then

$$C^*_B(I(B)) \cong A_B.$$

Furthermore, every unital AF algebra is isomorphic to the universal C*-algebra of a Boolean inverse monoid of the form $I(B)$ for some B.

Recall that a compact convex metrizable subset X of a locally convex space is a *Choquet simplex* if and only if for each $x \in X$ there exists a unique measure ν concentrated on the extreme points of X for which x is the center of gravity of X for ν [Phe01]. Now we can use the following seminal result of Blackadar to make a statement about the set of normalized invariant means for AF inverse monoids.

Theorem 5.2. (Blackadar, see [Bla80, Theorem 3.10]) Let Δ be any metrizable Choquet simplex. Then there exists a unital simple AF algebra A such that $T(A)$ is affinely isomorphic to Δ.

Corollary 5.3. Let Δ be any metrizable Choquet simplex. Then there exists an AF inverse monoid S such that $M(S)$ is affinely isomorphic to Δ.

Proof. This result follows from Theorem 4.13 because $\mathcal{G}_{\text{tight}}(S)$ is Hausdorff, amenable, and principal for every AF inverse monoid S. ☐

5.2 The 3×3 matrices

This example is a subexample of the previous example, but it will illustrate how we approach the following two examples.

Let $\mathcal{I}_3$ denote the symmetric inverse monoid on the three element set $\{1, 2, 3\}$. This is a Boolean inverse monoid which satisfies condition (H), and we define a map $\pi : \mathcal{I}_3 \to \mathbb{M}_3$ by saying that

$$\pi(\phi)_{ij} = \begin{cases} 1 & \text{if } \phi(j) = i \\ 0 & \text{otherwise.} \end{cases}$$

Then it is straightforward to verify that π is in fact the universal Boolean inverse monoid representation of $\mathcal{I}_3$.
Now instead consider the subset $R_3 \subset \mathcal{I}_3$ consisting of the identity, the empty function, and all functions with domain consisting of one element. Then R_3 is an inverse monoid, and $\pi(R_3)$ is the set of all matrix units together with the identity matrix and zero matrix. When restricted to R_3, π is the universal tight representation of R_3. Hence $C^*_{\text{tight}}(R_3) \cong C^*_B(\mathcal{I}_3) \cong \mathbb{M}_3$.
There is only one invariant mean μ on $\mathcal{I}_3$ – for an idempotent $\text{Id}_U \in \mathcal{I}_3$, we have $\mu(\text{Id}_U) = \frac{1}{3}|U|$. The tight groupoid of R_3 is the equivalence relation $\{1,2,3\} \times \{1,2,3\}$, which is principal – we also have that $\mathcal{G}_{\text{tight}}(R_3)^a \cong \mathcal{I}_3$. The unique invariant mean on $\mathcal{I}_3$ is identified with the unique normalized trace on $\mathbb{M}_3$.
Our last two examples follow this mold, where we have an inverse monoid S which generates a C*-algebra $C^*_{\text{tight}}(S)$, and we relate the traces of $C^*_{\text{tight}}(S)$ to the invariant means of $\mathcal{G}_{\text{tight}}(S)^a$.

5.3 Self-similar groups

Let X be a finite set, let G be a group, and let X^* denote the set of all words in elements of X, including an empty word $\varnothing$. Let X^ω denote the Cantor set of one-sided infinite words in X, with the product topology of the discrete topology on X. For $\alpha \in X^*$, let $C(\alpha) = \{\alpha x \mid x \in X^\omega\}$ – sets of this type are called *cylinder sets* and form a clopen basis for the topology on X.
Suppose that we have a faithful length-preserving action of G on X^*, with $(g, \alpha) \mapsto g \cdot \alpha$, such that for all $g \in G$, $x \in X$ there exists a unique element of G, denoted $g|_x$, such that for all $\alpha \in X^*$

$$g(x\alpha) = (g \cdot x)(g|_x \cdot \alpha).$$

In this case, the pair (G, X) is called a *self-similar group.* The map $G \times X \to G$, $(g, x) \mapsto g|_x$ is called the *restriction* and extends to $G \times X^*$ via the formula

$$g|_{\alpha_1 \cdots \alpha_n} = g|_{\alpha_1}|_{\alpha_2} \cdots |_{\alpha_n}$$

and this restriction has the property that for $\alpha, \beta \in X^*$, we have

$$g(\alpha\beta) = (g \cdot \alpha)(g|_\alpha \cdot \beta).$$

The action of G on X^* extends to an action of G on X^ω given by

$$g \cdot (x_1 x_2 x_3 \dots) = (g \cdot x_1)(g|_{x_1} \cdot x_2)(g|_{x_1 x_2} \cdot x_3) \cdots$$

In [Nek09], Nekrashevych associates a C*-algebra to (G, X), denoted $\mathcal{O}_{G,X}$, which is the universal C*-algebra generated by a set of isometries $\{s_x\}_{x \in X}$ and a unitary representation $\{u_g\}_{g \in G}$ satisfying

(i) $s_x^* s_y = 0$ if $x \neq y$,

(ii) $\sum_{x\in X} s_x s_x^* = 1$,

(iii) $u_g s_x = s_{g\cdot x} u_{g|_x}$.

One can also express $\mathcal{O}_{G,X}$ as the tight C*-algebra of an inverse semigroup. Let

$$S_{G,X} = \{(\alpha, g, \beta) \mid \alpha, \beta \in X^*, g \in G\} \cup \{0\}.$$

This set becomes an inverse semigroup when given the operation

$$(\alpha, g, \beta)(\gamma, h, \nu) = \begin{cases} (\alpha(g\cdot\gamma'), g|_{\gamma'} h, \nu), & \text{if } \gamma = \beta\gamma', \\ (\alpha, g(h^{-1}|_{\beta'})^{-1}, \nu(h^{-1}\cdot\beta')), & \text{if } \beta = \gamma\beta', \\ 0 & \text{otherwise} \end{cases}$$

with

$$(\alpha, g, \beta)^* = (\beta, g^{-1}, \alpha).$$

Here, $E(S_{X,G}) = \{(\alpha, 1_G, \alpha) \mid \alpha \in X^*\}$, and the tight spectrum $\widehat{E}_{\text{tight}}(S_{G,X})$ is homeomorphic X^ω by the identification

$$x \in X^\omega \mapsto \{(\alpha, 1_G, \alpha) \in E(S_{G,X}) \mid \alpha \text{ is a prefix of } x\} \in \widehat{E}_{\text{tight}}(S_{G,X}).$$

If θ is the standard action of $S_{G,X}$ on $\widehat{E}_{\text{tight}}(S_{G,X})$, then $D^\theta_{(\alpha,1_G,\alpha)} = C(\alpha)$. If $s = (\alpha, g, \beta) \in S_{X,G}$, then

$$\theta_s : C(\beta) \to C(\alpha)$$

$$\theta_s(\beta x) = \alpha(g\cdot x)$$

It is shown in [EP14] that $\mathcal{O}_{G,X}$ is isomorphic to $C^*_{\text{tight}}(S_{G,X})$.

We show that the universal tight representation of $S_{G,X}$ is faithful. This will be accomplished if we can show that the map from $S_{G,X}$ to $\mathcal{G}_{\text{tight}}(S_{G,X})^a$ given by

$$s \mapsto \Theta(s, D^\theta_{s^*s})$$

is injective. If $s = (\alpha, g, \beta)$, then

$$\Theta(s, D^\theta_{s^*s}) = \{[(\alpha, g, \beta), \beta x] \mid x \in X^\omega\}.$$

It is straightforward that $d(\Theta(s, D^\theta_{s^*s})) = C(\beta)$ and $r(\Theta(s, D^\theta_{s^*s})) = C(\alpha)$. Suppose we have another element $t = (\gamma, h, \eta)$ such that $\Theta(s, D^\theta_{s^*s}) = \Theta(t, D^\theta_{t^*t})$. Since these two bisections are equal, their sources (resp. ranges) must be equal, so $C(\beta) = C(\eta)$ (resp. $C(\alpha) = C(\gamma)$). Hence, $\alpha = \gamma$ and $\beta = \eta$. Since r and d are both bijective on these slices, we must have that for all $\beta x \in C(\beta)$, $\alpha(g\cdot x) = \alpha(h\cdot x)$. Hence for all $x \in X^\omega$, we must have that $g\cdot x = h\cdot x$. The action of G on X^* is faithful, so the induced action of G on X^ω is also faithful, hence $g = h$ and so $t = s$.

As it stands, the Boolean inverse monoid $\mathcal{G}_{\text{tight}}(S_{G,X})^a$ cannot have any invariant means. This is because the subalgebra of $\mathcal{O}_{G,X}$ generated by $\{s_x \mid x \in X\}$

is isomorphic to the Cuntz algebra $\mathcal{O}_{|X|}$, and a trace on $\mathcal{O}_{G,X}$ would have to restrict to a trace on $\mathcal{O}_{|X|}$, which is purely infinite and hence has no traces.

To justify the inclusion of this example in this paper about invariant means, we restrict to an inverse subsemigroup of $S_{G,X}$ whose corresponding ample semigroup will admit an invariant mean. Let

$$S_{G,X}^{=} = \{(\alpha, g, \beta) \in S_{G,X} \mid |\alpha| = |\beta|\} \cup \{0\}.$$

One can easily verify that this is closed under product and involution, and so is a inverse subsemigroup of $S_{G,X}$, with the same set of idempotents as $S_{G,X}$. If $\alpha, \beta \in X^*$, $|\alpha| = |\beta|$, and $g \in G$, then

$$(\alpha, g, \beta)^*(\alpha, g, \beta) = (\beta, 1_G, \beta), \qquad (\alpha, g, \beta)(\alpha, g, \beta)^* = (\alpha, 1_G, \alpha).$$

If μ were an invariant mean on $\mathcal{G}_{\mathrm{tight}}(S_{G,X}^{=})^a$, then we would have to have, for all $\alpha, \beta \in X^*$ and $|\alpha| = |\beta|$, that $\mu(C(\alpha)) = \mu(C(\beta))$. Moreover, for a given length n, the set $\{C(\alpha) \mid |\alpha| = n\}$ forms a disjoint partition of X^ω, and so we must have

$$\mu(C(\alpha)) = |X|^{-|\alpha|}. \tag{9}$$

Any clopen subset of X^ω must be a finite disjoint union of cylinders. Hence the map μ on $E(\mathcal{G}_{\mathrm{tight}}(S_{G,X}^{=})^a)$ determined by (9) is an invariant mean, and is in fact the unique invariant mean on $\mathcal{G}_{\mathrm{tight}}(S_{G,X}^{=})^a$.

In the general case, it is possible for $\mathcal{G}_{\mathrm{tight}}(S_{G,X}^{=})$ to be neither Hausdorff nor principal. We now give an explicit example where we get a unique trace to go along with our unique invariant mean.

Example 5.4. (The 2-odometer)

Let $X = \{0, 1\}$ and let $\mathbb{Z} = \langle z \rangle$ be the group of integers with identity e written multiplicatively. The *2-odometer* is the self-similar group $(\mathbb{Z}, X)$ determined by

$$z \cdot 0 = 1 \qquad z|_0 = e$$

$$z \cdot 1 = 0 \qquad z|_1 = z.$$

If one views a word $\alpha \in X^*$ as a binary number (written backwards), then $z \cdot \alpha$ is the same as 1 added to the binary number for α, truncated to the length of α if needed. If such truncation is not needed, $z|_\alpha = e$, but if truncation is needed, $z|_\alpha = z$.

The action of $\mathbb{Z}$ on $\{0, 1\}^\omega$ induced by the 2-odometer is the familiar Cantor minimal system of the same name. For $x \in \{0, 1\}^\omega$ we have

$$z \cdot x = \begin{cases} 000\cdots & \text{if } x_i = 1 \text{ for all } i \\ 00\cdots 01x_{i+1}x_{i+2}\cdots & \text{if } x_i = 0 \text{ and } x_j = 1 \text{ for all } j < i \end{cases}$$

This action of $\mathbb{Z}$ is *free* (i.e. $z^n \cdot x = x$ implies $n = 0$) and *minimal* (i.e. the set $\{z^n \cdot x \mid n \in \mathbb{Z}\}$ is dense in $\{0, 1\}^\omega$ for all $x \in \{0, 1\}^\omega$).

Lemma 5.5. The groupoid of germs $\mathcal{G}_{\text{tight}}(S^{=}_{\mathbb{Z},X})$ is principal.

Proof. Take $x, y \in \{0,1\}^\omega$ and suppose that we have $\alpha, \beta \in \{0,1\}^*$ with $|\alpha| = |\beta|$ and $n \in \mathbb{Z}$ such that $[(\alpha, z^n, \beta), x] \in \mathcal{G}_{\text{tight}}(S^{=}_{\mathbb{Z},X})$ and $r([(\alpha, z^n, \beta), x]) = y$. This implies that $x = \beta v$ for some $v \in \{0,1\}^\omega$, and that $y = \alpha(z^n \cdot v)$. Suppose we can find another germ from x to y, that is, suppose we have $\gamma, \eta \in \{0,1\}^*$ with $|\gamma| = |\eta|$ and $m \in \mathbb{Z}$ such that $[(\gamma, z^n, \eta), x] \in \mathcal{G}_{\text{tight}}(S_{\mathbb{Z},X})$ and $r([(\gamma, z^n, \eta), x]) = y$. Again we can conclude that $x = \eta u$ for some $u \in \{0,1\}^\omega$, and that $y = \gamma(z^m \cdot u)$. There are two cases.
Suppose first that $\beta = \eta\delta$ for some $\delta \in \{0,1\}^*$. Then $\eta\delta v = x = \eta u$, and so $\delta v = u$. We also have $\alpha(z^n \cdot v) = y = \gamma(z^m \cdot u)$. Because $|\alpha| = |\beta| \geq |\eta| = |\gamma|$, this implies that there exists $\nu \in \{0,1\}^*$ with $|\nu| = |\delta|$ and $\alpha = \gamma\nu$. Hence $\nu(z^n \cdot v) = z^m \cdot u = (z^m \cdot \delta)\, z^m|_\delta \cdot v$, which gives us that $\nu = z^m \cdot \delta$ and $z^n \cdot v = z^m|_\delta \cdot v$, and since the action on $\{0,1\}^\omega$ is free we have $z^n = z^m|_\delta$. So we have that $x \in C(\beta) = D^\theta_{(\beta,e,\beta)}$, and we calculate

$$\begin{aligned}(\gamma, z^n, \eta)(\beta, e, \beta) &= (\gamma(z^m \cdot \delta),\, z^m|_\delta\,, \beta) = (\gamma\nu, z^n, \beta) = (\alpha, z^n, \beta)\\ &= (\alpha, z^n, \beta)(\beta, e, \beta)\end{aligned}$$

where the first equality is by the definition of the product. Hence $[(\alpha, z^n, \beta), x] = [(\gamma, z^n, \eta), x]$. The case where β is shorter than η is similar. Hence, $\mathcal{G}_{\text{tight}}(S^{=}_{\mathbb{Z},X})$ is principal. ☐

It is routine to check that $S^{=}_{\mathbb{Z},X}$ satisfies condition (H) (in fact, it is E*-unitary, see [ES16, Example 3.4]). The groupoid $\mathcal{G}_{\text{tight}}(S^{=}_{\mathbb{Z},X})$ is amenable, see [ADR00, Proposition 5.1.1] and [EP13, Corollary 10.18]. Hence Theorem 4.13 applies, and there is only one normalized trace on $C^*_{\text{tight}}(S^{=}_{\mathbb{Z},X})$, the one arising from the invariant mean.
As the observant reader is no doubt aware at this point, $C^*_{\text{tight}}(S^{=}_{\mathbb{Z},X})$ is nothing more than the crossed product $C(\{0,1\}^\omega) \rtimes \mathbb{Z}$ arising from the usual odometer action [Nek04, Theorem 7.2], which has a unique normalized trace due to the dynamical system $(\{0,1\}^\omega, \mathbb{Z})$ having a unique invariant measure (given by (9)).

5.4 Aperiodic tilings

We close with another example where the traces on the relevant C*-algebras are known beforehand, and hence give us invariant means.
A *tile* is a closed subset of $\mathbb{R}^d$ homeomorphic to the closed unit ball. A *partial tiling* is a collection of tiles in $\mathbb{R}^d$ with pairwise disjoint interiors, and the *support* of a partial tiling is the union of its tiles. A *patch* is a finite partial tiling, and a *tiling* is a partial tiling with support equal to $\mathbb{R}^d$. If P is a partial tiling and $U \subset \mathbb{R}^d$, then let $P(U)$ be the partial tiling of all tiles in P which intersect U. A tiling T is called *aperiodic* if $T + x \neq T$ for all $0 \neq x \in \mathbb{R}^d$.
Let T be a tiling. We form an inverse semigroup S_T from T as follows. For a patch $P \subset T$ and tiles $t_1, t_2 \in P$ we call the triple (t_1, P, t_2) a *doubly pointed patch*. We put an equivalence relation on such triples, by saying that $(t_1, P, t_2) \sim (r_1, Q, r_2)$ if there exists a vector $x \in \mathbb{R}^d$ such that

$(t_1 + x, P + x, t_2 + x) = (r_1, Q, r_2)$, and let $[t_1, P, t_2]$ denote the equivalence class of such a triple – this is referred to a *doubly pointed patch class*. Let

$$S_T = \{[t_1, P, t_2] \mid (t_1, P, t_2) \text{ is doubly pointed patch } \} \cup \{0\}$$

be the set of all doubly pointed patch classes together with a zero element. If $[t_1, P, t_2], [r_1, Q, r_2]$ are two elements of S_T, we let

$$[t_1, P, t_2][r_1, Q, r_2] = \begin{cases} [t_1, P \cup Q', r_2'] & \text{if there exists } (r_1', Q', r_2') \in [r_1, Q, r_2] \\ & \text{such that } r_1' = t_2 \text{ and } P \cup Q' \text{ is} \\ & \text{a patch in } T + x \text{ for some } x \in \mathbb{R}^d \\ 0 & \text{otherwise,} \end{cases}$$

and define all products involving 0 to be 0. Also, let $[t_1, P, t_2]^* = [t_2, P, t_1]$. With these operations, S_T is an inverse semigroup. This inverse semigroup was defined by Kellendonk [Kel97a] [Kel97b], and is E*-unitary.

Suppose there exists a finite set $\mathcal{P}$ of tiles each of which contain the origin in the interior such that for all $t \in T$, there exists $x_t \in \mathbb{R}^d$ and $p \in \mathcal{P}$ such that $t = p + x_t$. In this case, $\mathcal{P}$ is called a set of *prototiles* for T. By possibly adding labels, we may assume that x_t and p are unique – we call x_t the *puncture* of t. Consider the set

$$X_T = \{T - x_t \mid t \in T\}$$

and put a metric on X_T by setting

$$d(T_1, T_2) = \inf\{1, \epsilon \mid T_1(B_{1/\epsilon}(0)) = T_1(B_{1/\epsilon}(0))\}$$

and let Ω_{punc} denote the completion of X_T in this metric (above, $B_r(x)$ denotes the open ball in $\mathbb{R}^d$ of radius r around $x \in \mathbb{R}$). One can show that all elements of Ω_{punc} are tilings consisting of translates of $\mathcal{P}$ which also contain an element of $\mathcal{P}$ and that the metric above extends to the same metric on Ω_{punc} – this is called the *punctured hull* of T.

We make the following assumptions on T:

1. T has *finite local complexity* if for any $r > 0$, there are only finitely many patches in T with supports having outer radius less than r, up to translational equivalence.

2. T is *repetitive* if for every patch $P \subset T$, there exists $R > 0$ such that every ball of radius R in $\mathbb{R}^d$ contains a translate of P.

3. T is *strongly aperiodic* if all elements of Ω_{punc} are aperiodic.

In this case Ω_{punc} is homeomorphic to the Cantor set. For a patch $P \subset T$ and tile $t \in P$, let

$$U(P, t) = \{T' \in \Omega_{\text{punc}} \mid P - x_t \subset T'\}.$$

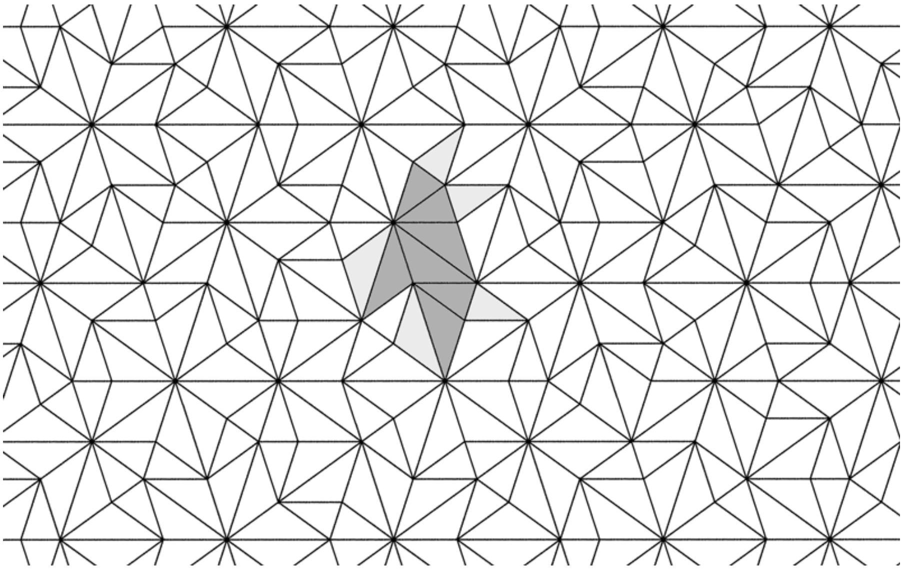


Figure 1: In the Robinson triangles version of the Penrose tiling, each triangle is always next to a similar triangle with which it forms a rhombus. Let P be the dark gray patch, and let P' be the patch with the lighter gray tiles added. Then for any dark gray tile t, $U(P,t) = U(P',t)$

Then these sets are clopen in Ω_{punc} and generate the topology. Let

$$\mathcal{R}_{\text{punc}} = \{(T_1, T_1 + x) \in \Omega_{\text{punc}} \times \Omega_{\text{punc}} \mid x \in \mathbb{R}^d\}$$

and view this equivalence relation as a principal groupoid. Endow it with the topology inherited by viewing it as a subspace of $\Omega_{\text{punc}} \times \mathbb{R}^d$. For a patch $P \subset T$ and $t_1, t_2 \in P$, let

$$V(t_1, P, t_2) = \{(T_1, T_2) \in \mathcal{R}_{\text{punc}} \mid T_1 \in U(P, t_1), T_2 = T_1 + x_{t_1} - x_{t_2}\}$$

Then these sets are compact bisections in $\mathcal{R}_{\text{punc}}$, and generate the topology on $\mathcal{R}_{\text{punc}}$. This groupoid is Hausdorff, ample, and amenable [PS99]. The C*-algebra of $\mathcal{R}_{\text{punc}}$ was defined by Kellendonk in [Kel95] (denoted there A_T) and studied further in [KP00], [Put00], [Put10], [Phi05], [Sta14].

We proved in [EGS12, Theorem 3] that $\mathcal{G}_{\text{tight}}(S_T) \cong \mathcal{R}_{\text{punc}}$ – the universal tight representation of S_T maps $[t_1, P, t_2]$ to the characteristic function of $V(t_1, P, t_2)$.[4] It is interesting to note that in this case that the universal tight representation may not be faithful. Suppose that we could find $P \subset P'$, both patches in T, and that $P + x \subset T$ can only happen if $P' + x \subset T$. Then

[4]The same result follows from [Len08, Section 9] combined with the fact that Lenz's groupoid coincides with the tight groupoid when $\widehat{E}_{\text{tight}}(S) = \widehat{E}_\infty(S)$, see [LL13, Theorem 5.15].

for a tile $t \in P$, the two idempotents $[t, P, t], [t, P', t]$ are different elements in S_T, but are both mapped to the characteristic function of $U(P, t) = U(P', t)$ under the universal tight representation – indeed, $[t, P', t]$ is dense in $[t, P, t]$, see Figure 1. We note that [Len08], [Exe09], and [LL13] address other cases where the tight representation may not be faithful.
The C*-algebra A_T can be seen as the C*-algebra of a Boolean inverse monoid, namely $\mathcal{G}_{\text{tight}}(S_T)^a$ – one could then rightly call this the *Boolean inverse monoid associated to* T. The traces of A_T are already well-studied, see [KP00], [Put00]. Often, as is the case with the Penrose tiling, there is a unique trace, see [Put00].

Theorem 5.6. Let T be a tiling which satisfies conditions 1–3 above, and let $\mathcal{G}_{\text{tight}}(S_T)^a$ be the Boolean inverse monoid associated to T. Then $M(\mathcal{G}_{\text{tight}}(S_T)^a) \cong T(A_T) \cong IM(\mathcal{R}_{\text{punc}})$.

Acknowledgment: I thank Thierry Giordano for an enlightening conversation about this work, and I also thank Ganna Kudryavtseva, Ruy Exel, and Daniel Lenz for comments and suggestions on a preprint version of this paper. Finally, I thank the referee for a thorough reading.

References

[AD16a] Claire Anantharaman-Delaroche. Exact Groupoids. *arXiv:1605.05117*, May 2016.

[AD16b] Claire Anantharaman-Delaroche. Some remarks about the weak containment property for groupoids and semigroups. *arXiv:1604.01724*, April 2016.

[ADR00] Claire Anantharaman-Delaroche and Jean Renault. *Amenable Groupoids.* Monographie de l'Enseignement mathématique. L'Enseignement Mathématique, 2000.

[BCFS14] Jonathan Brown, Lisa Orloff Clark, Cynthia Farthing, and Aidan Sims. Simplicity of algebras associated to étale groupoids. *Semigroup Forum*, 88(2):433–452, 2014.

[Bla80] Bruce Blackadar. Traces on simple AF C*-algebras. *J. Funct. Anal.*, 38(2):156 – 168, 1980.

[BO08] Nathanial P. Brown and Narutaka Ozawa. *C*-algebras and Finite-dimensional Approximations.* Graduate studies in mathematics. American Mathematical Soc., 2008.

[COP15] Toke Meier Carlsen, Enrique Ortega, and Eduardo Pardo. *C**-algebras associated to Boolean dynamical systems. *arXiv:1510.06718*, October 2015.

[DM14] Allan P. Donsig and David Milan. Joins and covers in inverse semigroups and tight C*-algebras. *B. Aust. Math. Soc.*, 90:121–133, 8 2014.

[EGS12] Ruy Exel, Daniel Gonçalves, and Charles Starling. The tiling C*-algebra viewed as a tight inverse semigroup algebra. *Semigroup Forum*, 84:229–240, 2012.

[EP13] Ruy Exel and Enrique Pardo. Graphs, groups and self-similarity. *arXiv:1307.1120*, July 2013.

[EP14] Ruy Exel and Enrique Pardo. Self-similar graphs, a unified treatment of Katsura and Nekrashevych C*-algebras. *arXiv:1409.1107*, September 2014.

[EP16] Ruy Exel and Enrique Pardo. The tight groupoid of an inverse semigroup. *Semigroup Forum*, 92:274 – 303, 2016.

[ER06] Ruy Exel and Jean Renault. AF-algebras and the tail-equivalence relation on Bratteli diagrams. *Proc. Amer. Math. Soc.*, 134(1):193–206, 2006.

[ES15] Ruy Exel and Charles Starling. Amenable actions of inverse semigroups. *Ergodic Theory Dyn. Syst.*, to appear 2015. doi:10.1017/etds.2015.60.

[ES16] Ruy Exel and Charles Starling. Self-similar graph C*-algebras and partial crossed products. *J. Operator Theory*, 75, 2016.

[Exe08] Ruy Exel. Inverse semigroups and combinatorial C^*-algebras. *Bull. Braz. Math. Soc. (N.S.)*, 39(2):191–313, 2008.

[Exe09] Ruy Exel. Tight representations of semilattices and inverse semigroups. *Semigroup Forum*, 79(1):159–182, 2009.

[Exe10] Ruy Exel. Reconstructing a totally disconnected groupoid from its ample semigroup. *Proc. Amer. Math. Soc.*, 138(8):2991–3001, 2010.

[GH09] Steven Givant and Paul Halmos. *Introduction to Boolean algebras.* Undergraduate Texts in Mathematics. Springer, 2009.

[Haa91] Uffe Haagerup. Quasitraces on exact C*-algebras are traces. *arXiv:1307.1120*, March 1991.

[Kel95] Johannes Kellendonk. Noncommutative geometry of tilings and gap labelling. *Rev. Math. Phys.*, 7(7):1133–1180, 1995.

[Kel97a] Johannes Kellendonk. The local structure of tilings and their integer group of coinvariants. *Commun. Math. Phys.*, 187(1):115–157, 1997.

[Kel97b] Johannes Kellendonk. Topological equivalence of tilings. *J. Math. Phys.*, 38(4), 1997.

[KL14] Ganna Kudryavtseva and Mark V. Lawson. A perspective on non-commutative frame theory. *arXiv:1404.6516*, April 2014.

[KLLR16] Ganna Kudryavtseva, Mark V. Lawson, Daniel H. Lenz, and Pedro Resende. Invariant means on Boolean inverse monoids. *Semigroup Forum*, 92(1):77–101, 2016.

[KP00] Johannes Kellendonk and Ian F. Putnam. Tilings, C^*-algebras, and K-theory. In *Directions in mathematical quasicrystals*, volume 13 of *CRM Monogr. Ser.*, pages 177–206. Amer. Math. Soc., Providence, RI, 2000.

[KR06] Alex Kumjian and Jean Renault. KMS states on C*-algebras associated to expansive maps. *Proc. Amer. Math. Soc.*, 134(7):2067–2078, 2006.

[Len08] Daniel H. Lenz. On an order-based construction of a topological groupoid from an inverse semigroup. *Proc. Edinb. Math. Soc. (2)*, 51(2):387–406, 2008.

[Li16] Xin Li. Partial transformation groupoids attached to graphs and semigroups. *arXiv:1603.09165*, March 2016.

[LL13] Mark V. Lawson and Daniel H. Lenz. Pseudogroups and their étale groupoids. *Adv. Math.*, 244(0):117 – 170, 2013.

[LLR00] Flemming Larsen, Niels Jacob Lausten, and Mikael Rørdam. *An Introduction to K-Theory for C*-Algebras.* London Mathematical Society Student Texts. Cambridge University Press, 2000.

[LS14] Mark V. Lawson and Phil Scott. AF inverse monoids and the structure of countable MV-algebras. *arXiv:1408.1231*, 2014.

[Nek04] Volodymyr Nekrashevych. Cuntz-Pimsner algebras of group actions. *J. Operator Theory*, 52:223–249, 2004.

[Nek09] Volodymyr Nekrashevych. C*-algebras and self-similar groups. *J. reine angew. Math*, 630:59–123, 2009.

[Pat99] Alan Paterson. *Groupoids, inverse semigroups, and their operator algebras.* Birkhäuser, 1999.

[Phe01] Robert Phelps. *Lectures on Choquet's Theorem.* Lecture notes in mathematics. Springer-Verlag Berlin Heidelberg, 2001.

[Phi05] N. Christopher Phillips. Crossed products of the Cantor set by free minimal actions of $\mathbb{Z}^d$. *Commun. Math. Phys.*, 256(1):1–42, 2005.

[PS99] Ian F. Putnam and Jack Spielberg. The structure of C*-algebras associated with hyperbolic dynamical systems. *J. Funct. Anal.*, 163(2):279 – 299, 1999.

[Put] Ian F. Putnam. Lecture Notes on C*-algebras. http://www.math.uvic.ca/faculty/putnam/ln/C*-algebras.pdf.

[Put00] Ian F. Putnam. The ordered K-theory of C^*-algebras associated with substitution tilings. *Commun. Math. Phys.*, 214(3):593–605, 2000.

[Put10] Ian F. Putnam. Non-commutative methods for the K-theory of C^*-algebras of aperiodic patterns from cut-and-project systems. *Commun. Math. Phys.*, 294(3):703–729, 2010.

[Ren80] Jean Renault. *A groupoid approach to C^*-algebras*, volume 793 of *Lecture Notes in Mathematics*. Springer, Berlin, 1980.

[Sta14] Charles Starling. Finite symmetry group actions on substitution tiling C*-algebras. *Münster J. of Math*, 7:381–412, 2014.

[Sta15] Charles Starling. Boundary quotients of C*-algebras of right LCM semigroups. *J. Funct. Anal.*, 268(11):3326 – 3356, 2015.

[Sta16] Charles Starling. Inverse semigroups associated to subshifts. *J. Algebra*, 463:211 – 233, 2016.

[Ste10] Benjamin Steinberg. A groupoid approach to discrete inverse semigroup algebras. *Adv. Math.*, 223(2):689 – 727, 2010.

[Ste16] Benjamin Steinberg. Simplicity, primitivity and semiprimitivity of étale groupoid algebras with applications to inverse semigroup algebras. *J. Pure Appl. Algebr.*, 220(3):1035 – 1054, 2016.

[Wil15] Rufus Willett. A non-amenable groupoid whose maximal and reduced C*-algebras are the same. *To appear in Münster J. of Math* arXiv:1504.05615 [math.OA], 2015.

Charles Starling
University of Ottawa
Department of Mathematics
and Statistics
585 King Edward
Ottawa, ON, Canada
K1N 6N5
cstar050@uottawa.ca

Operator-Valued Monotone Convolution Semigroups and an Extension of the Bercovici-Pata Bijection.

Michael Anshelevich[1], John D. Williams

Received: May 19, 2015
Revised: May 17, 2016

Communicated by Friedrich Götze

Abstract. In a 1999 paper, Bercovici and Pata showed that a natural bijection between the classically, free and Boolean infinitely divisible measures held at the level of limit theorems of triangular arrays. This result was extended to include monotone convolution by the authors in [AW14]. In recent years, operator-valued versions of free, Boolean and monotone probability have also been developed. Belinschi, Popa and Vinnikov showed that the Bercovici-Pata bijection holds for the operator-valued versions of free and Boolean probability. In this article, we extend the bijection to include monotone probability theory even in the operator-valued case. To prove this result, we develop the general theory of composition semigroups of non-commutative functions and largely recapture Berkson and Porta's classical results on composition semigroups of complex functions in operator-valued setting. As a byproduct, we deduce that operator-valued monotonically infinitely divisible distributions belong to monotone convolution semigroups. Finally, in the appendix, we extend the result of the second author on the classification of Cauchy transforms for non-commutative distributions to the Cauchy transforms associated to more general completely positive maps.

2010 Mathematics Subject Classification: Primary 46L53; Secondary 30D05, 47D06, 60E07

[1]The first author was supported in part by NSF grant DMS-1160849.

1. Introduction

It is a remarkable fact that there are natural bijections between the classes of infinitely divisible measures in each of the four universal non-commutative probability theories, which not only arise from the Lévy-Hinčin representations of the measures, but are maintained at the level of limit theorems of triangular arrays. This is made precise in the following theorem:

Theorem 1.1. Fix a finite positive Borel measure σ on $\mathbb{R}$, a real number γ, a sequence of probability measures $\{\mu_n\}_{n\in\mathbb{N}}$, and a sequence of positive integers $k_1 < k_2 < \cdots$. The following assertions are equivalent:

(a) (Classical / tensor) The sequence $\underbrace{\mu_n * \mu_n * \cdots * \mu_n}_{k_n}$ converges weakly to $\nu_*^{\gamma,\sigma}$;

(b) (Free) The sequence $\underbrace{\mu_n \boxplus \mu_n \boxplus \cdots \boxplus \mu_n}_{k_n}$ converges weakly to $\nu_\boxplus^{\gamma,\sigma}$;

(c) (Boolean) The sequence $\underbrace{\mu_n \uplus \mu_n \uplus \cdots \uplus \mu_n}_{k_n}$ converges weakly to $\nu_\uplus^{\gamma,\sigma}$;

(d) (Monotone) The sequence $\underbrace{\mu_n \rhd \mu_n \rhd \cdots \rhd \mu_n}_{k_n}$ converges weakly to $\nu_\rhd^{\gamma,\sigma}$;

(e) The measures

$$k_n \frac{x^2}{x^2+1} d\mu_n(x) \to \sigma$$

weakly, and

$$\lim_{n\uparrow\infty} k_n \int_{\mathbb{R}} \frac{x}{x^2+1} d\mu_n(x) = \gamma.$$

Here $\nu_*^{\gamma,\sigma}$, $\nu_\boxplus^{\gamma,\sigma}$, $\nu_\uplus^{\gamma,\sigma}$, $\nu_\rhd^{\gamma,\sigma}$ are probability measures defined explicitly through their complex-analytic transforms. The equivalence of classical, free, and Boolean limit theorems in parts (a), (b), (c) and (e) was proven in a by now classic paper due to Bercovici and Pata [BP99]. The monotone non-commutative probability theory is of more recent provenance [Mur00, Mur01]. The inclusion of part (d) was proven in our recent paper [AW14].

Voiculescu developed operator-valued notions of non-commutative probability [Voi87] where probability measures are replaced by certain completely positive maps from the ring of non-commutative polynomials over a C*-algebra. An analogous theorem in this more general setting, namely the equivalence of parts (b) and (c), was proven in [BPV12]. The first main result in this paper is the inclusion of (d) at this level of generality.

In order to study monotone infinitely divisible $\mathcal{B}$-valued distributions, we must first develop the theory of composition semigroups of non-commutative functions in a manner analogous to Berkson and Porta's study of these semigroups at the level of complex functions [BPo78]. This stems from the fact that the convolution operation for monotone probability theory satisfies the following relation for the associated F-transforms,

$$F_{\mu \rhd \nu} = F_\mu \circ F_\nu,$$

so that infinitely-divisible distributions form such a composition semigroup. In the second main result of the paper, we prove that any monotone infinitely-divisible distribution can be included in such a semigroup. Note that even in the scalar-valued case, this is a recent result, proved by Serban Belinschi in his thesis. Finally, we characterize generators of such composition semigroups, and a smaller set of generators of composition semigroups of F-transforms.

In Section 2, we provide background and preliminary results. In section 3, we study composition semigroups of vector-valued and non-commutative analytic functions. The main results of this section are Proposition 3.3, which shows that there is a natural notion of a time derivative for semigroups of vector-valued analytic functions $\{f_t\}_{t\geq 0}$, and Theorem 3.5, which proves that, in the case of F-transforms and more general self-maps of the complex upper half plane, these semi-groups are in bijection with certain classes of functions defined through their analytic and asymptotic properties. This bijection provides a Lévy-Hinčin representation for these infinitely divisible distributions. In section 4 we prove the main result of the paper, namely the extension of Theorem 1.1 to the operator-valued case. In contrast to the previous section, this is achieved through a combinatorial methodology. We close the paper with the Appendix, which is primarily concerned with the extension of the main result in [Wil13], namely the classification of the Cauchy transforms associated to $\mathcal{B}$-valued distributions, to a more general class of functions including the Cauchy transforms associated to more general CP maps.

Acknowledgements. We are grateful to the referee for helpful comments.

2. Preliminaries

Let $\mathcal{B}$ denote a unital C*-algebra and X a self-adjoint symbol. We will define the ring of noncommutative polynomials $\mathcal{B}\langle X\rangle$ as the algebraic free product of $\mathcal{B}$ and X. $\mathcal{B}_0\langle X\rangle$ are polynomials in $\mathcal{B}\langle X\rangle$ with zero constant term.

Definition 2.1. Let $\mu : \mathcal{B}\langle X\rangle \to \mathcal{B}$ denote a linear map. We say that μ is *exponentially bounded* with constant M if

$$\|\mu(b_1Xb_2\cdots Xb_{n+1})\| \leq M^n\|b_1\|\|b_2\|\cdots\|b_{n+1}\| \tag{1}$$

We abuse terminology and say that the map μ is *completely positive* (CP) if

$$(\mu\otimes 1_n)\left(\left[P_i(X)P_j^*(X)\right]_{i,j=1}^n\right) \geq 0 \tag{2}$$

for every family $P_i(X) \in \mathcal{B}\langle X\rangle$.

We define a set Σ_0 to be those $\mathcal{B}$-bimodular linear maps μ satisfying (1) and (2).

For a general introduction to non-commutative functions, we refer to [KVV14]. Throughout, $\mathcal{B}, \mathcal{A}$ shall denote unital C*-algebras. Let $M_n(\mathcal{B})$ denote the $n\times n$ matrices with entries in $\mathcal{B}$. We define the *noncommutative space over* $\mathcal{B}$ to be the set $\mathcal{B}_{nc} = \{M_n(\mathcal{B})\}_{n=1}^{\infty}$. A *non-commutative set* is a subset $\Omega \subset \mathcal{B}_{nc}$ that respects direct sums. That is, for $X \in \Omega\cap M_n(\mathcal{B})$ and $Y \in \Omega\cap M_p(\mathcal{B})$ we have that $X\oplus Y \in \Omega\cap M_{n+p}(\mathcal{B})$. We note that these definitions apply to the more general case of $\mathcal{B}$ being any unital, commutative ring, but we focus on the C^*-algebraic setting. Given $b \in M_n(\mathcal{B})$, the *non-commutative ball* of radius δ about b is the set $B_\delta^{nc}(b) := \sqcup_{k=1}^\infty B_\delta(\oplus^k b)$ where $B_\delta(\oplus^k b) \subset M_{nk}(\mathcal{B})$ is the standard ball of radius δ. A *non-commutative* function is a map $f : \Omega \to \mathcal{A}_{nc}$ with the following properties:

(a) $f(\Omega_n) \subset M_n(\mathcal{A})$
(b) f respects direct sums : $f(X \oplus Y) = f(X) \oplus f(Y)$
(c) f respects similarities: For $X \in \Omega_n$ and $S \in M_n(\mathbb{C})$ invertible we have that

$$f(SXS^{-1}) = Sf(X)S^{-1}$$

provided that $SXS^{-1} \in \Omega_n$.

A non-commutative function is said to be *locally bounded in slices* if, for every n and element $x \in \Omega_n$, $f|_{\Omega_n}$ is bounded on some neighborhood of x in the norm topology. It is a remarkable fact originally due to Taylor ([Tay72], [Tay73]) that a non-commutative function that is Gâteaux differentiable and locally bounded in slices is in fact analytic. A non-commutative function is *uniformly analytic* at $b \in M_n(\mathcal{B})$ if it is analytic and bounded on $B_r^{nc}(b)$ for some $r > 0$.

Let $M_n^{+,\epsilon}(\mathcal{B}) \subset M_n(\mathcal{B})$ denote those element $b \in M_n(\mathcal{B})$ with $\Im(b) > \epsilon 1_n$ and $M_n^+(\mathcal{B}) = \cup_{\epsilon>0} M_n^{+,\epsilon}$. We form a non-commutative set

$$H^+(\mathcal{B}) = \sqcup_{n=1}^\infty M_n^+(\mathcal{B})$$

and refer to this set as the *non-commutative upper half plane*.

We define a family of sets in $H^+(\mathcal{B})$. For $\alpha, \epsilon > 0$ define a *non-commutative Stolz angle* to be

$$\Gamma_{\alpha,\epsilon}^{(n)} := \{b \in M_n^{+,\epsilon}(\mathcal{B}) : \Im(b) > \alpha\Re(b)\}.$$

Let $\mu \in \Sigma_0$. We define the *Cauchy transform* of μ to be the analytic, non-commutative function $G_\mu = \{G_\mu^{(n)}\}_{n=1}^\infty$ such that

$$G_\mu^{(n)}(b) := (\mu \otimes 1_n)((b - X \otimes 1_n)^{-1}) : H^+(\mathcal{B}) \mapsto H^-(\mathcal{B}).$$

From this map, we may construct the *moment generating function*, the *F-transform*, the *Voiculescu transform* and the *$\mathcal{R}$-transform* respectively through the following equalities:

$$H^{(n)}(b) := G^{(n)}(b^{-1}) : H^-(\mathcal{B}) \mapsto H^-(\mathcal{B})$$
$$F^{(n)}(b) := G^{(n)}(b)^{-1} : H^+(\mathcal{B}) \mapsto H^+(\mathcal{B})$$
$$\varphi_\mu^{(n)}(b) := (F_\mu^{(n)})^{\langle -1 \rangle}(b) - b$$
$$\mathcal{R}_\mu^{(n)}(b) := \varphi_\mu^{(n)}(b^{-1})$$

where the superscript $\langle -1 \rangle$ refers to the composition inverse. We also note that the moment generating function extend to a neighborhood of 0 for $\mu \in \Sigma_0$ and that the Voiculescu-transform is only defined on a subset of $H^+(\mathcal{B})$. The following result, proven in [Wil13] and [PV13], classifies the F-transforms in terms of their analytic and asymptotic properties.

Theorem 2.1. Let $f = (f^{(n)}) : H^+(\mathcal{B}) \to H^+(\mathcal{B})$ denote an analytic, noncommutative function. The following conditions are equivalent.

(a) $f = F_\mu$ for some $\mu \in \Sigma_0$.

(b) The noncommutative function $k = (k^{(n)})_{n=1}^{\infty}$ defined by $k^{(n)}(b) := (f^{(n)}(b^{-1}))^{-1}$ has uniformly analytic extension to a neighborhood of 0. Moreover, for any sequence $\{b_k\}_{k\in\mathbb{N}}$ with $\|b_k^{-1}\| \downarrow 0$, $b_k^{-1} f^{(n)}(b_k) \to 1_n$ in norm.

(c) There exists an $\alpha \in \mathcal{B}$ and a $\sigma : \mathcal{B}\langle X\rangle \to \mathcal{B}$ which satisfies (1) and (2) such that, for all $n \in \mathbb{N}$,

$$f^{(n)}(b) = \alpha 1_n + b - (\sigma \otimes 1_n)(b(1 - Xb)^{-1}).$$

Moreover, the map σ in (c) is of the form $\sigma(P(X)) = \rho(XP(X)X)$ for ρ such that its restriction to $\mathcal{B}_0\langle X\rangle$ is positive.

We will require several classical results in complex function theory to prove our results. Theorem 3.16.3 in [HP74] is a useful analogue of the classical Cauchy estimates in complex analysis. We also refer to this reference for an overview of the differential structure of vector valued functions, including the higher order derivative δ^n utilized below.

Theorem 2.2. Let f be Gâteaux differentiable in $\mathcal{U}$ and assume that $\|f(x)\| \leq M$ for $x \in \mathcal{U}$. Then

$$\|\delta^n f(a; h)\| \leq Mn!$$

for $a + h \in \mathcal{U}$.

Further, theorem 3.17.17 in [HP74] provides Lipschitz estimates for analytic functions. Indeed, for an analytic function f that is locally bounded by $M(a)$ in a neighborhood of radius r_a, we have that

$$\|f(y) - f(x)\| \leq \frac{2M(a)\|x - y\|}{r_a - 2\|x - y\|} \tag{3}$$

Notation 2.2. We define a family Λ of functions $\Phi : H^+(\mathcal{B}) \to \overline{H^-(\mathcal{B})}$ through the following properties:

(i) The map $\mathcal{R}(b) := \Phi(b^{-1})$ has uniformly analytic continuation to a noncommutative ball about 0 with $\mathcal{R}(b)^* = \mathcal{R}(b)$

(ii) For any sequence $\{b_k\}_{k\in\mathbb{N}} \in \mathcal{B}$ with $\|b_k^{-1}\| \downarrow 0$, we have that $b_k^{-1}\Phi(b_k) \to 0$.

We also define a larger family of functions $\tilde{\Lambda}$ by replacing (i) and (ii) with the following weaker conditions

(I) For any $\epsilon > 0$, Φ is uniformly bounded on $\sqcup_{n=1}^{\infty} M_n^{+,\epsilon}(\mathcal{B})$.

(II) For any $\alpha, \epsilon > 0$ and a sequence $\{b_k\}_{k\in\mathbb{N}} \in \Gamma_{\alpha,\epsilon}^{(n)}$ with $\|b_k^{-1}\| \downarrow 0$, we have that $b_k^{-1}\Phi(b_k) \to 0$.

Definition 2.3. Let $\mu, \nu \in \Sigma_0$. We define the *monotone convolution* to be the noncommutative operation $(\mu, \nu) \mapsto \mu \rhd \nu \in \Sigma_0$ defined implicitly though the equality

$$F_{\mu\rhd\nu} := F_\mu \circ F_\nu.$$

Note that this definition uses Theorem 2.1 in an essential way, to show that a composition of F-transforms is an F-transform. See Section 4 and references [Pop08, HS11, Pop12, HS14] for the relation between this definition and monotone independence of Muraki.

DEFINITION 2.4. We say that μ is a $\triangleright$-infinitely divisible distribution if, for every n, there exists a distribution $\mu_n \in \Sigma_0$ such that

$$\mu = \underbrace{\mu_n \triangleright \mu_n \triangleright \cdots \triangleright \mu_n}_{n\ times} \tag{4}$$

We define a composition semigroup of F-transforms $\{F_t\}_{t\in\mathbb{Q}^+}$ by letting $F_{p/q} := F_{\mu_q}^{\circ p}$ where $\mu = \mu_q^{\triangleright q}$ for all $p, q \in \mathbb{N}$. We will show in Theorem 3.5 that this semigroup extends to an $\mathbb{R}^+$ semigroup, which moreover is generated by a function $\Phi \in \Lambda$ in a sense that will be made specific. Moreover, one of the main results in [Wil13] is that the set Λ is exactly the set of Voiculescu transforms associated to $\boxplus$-infinitely divisible distributions. This is not a coincidence and will drive the main result of this paper.

3. LÉVY-HINČIN REPRESENTATIONS FOR SEMIGROUPS OF NON-COMMUTATIVE FUNCTIONS.

We begin this section with a result showing that the divisors of $\triangleright$-infinitely divisible distributions maintain the same exponential bound. A similar result can be proven in the combinatorial setting of Section 4 in an easier manner, but the bound is less sharp.

PROPOSITION 3.1. Let μ denote a $\triangleright$-infinitely divisible distribution with exponential bound M. Then, for each k, the distribution μ_k satisfying $\mu = \mu_k^{\triangleright k}$ has exponential bound M.

Proof. Let $Xb_1Xb_2\cdots b_{n-1}X = Q(X) \in \mathcal{B}\langle X\rangle$ such that $\|b_1\| = \|b_2\| = \cdots\|b_{n-1}\| = 1$ and assume, for the sake of contradiction, that $\|\mu_k(Q(X))\| > M^n$. Then, using the Schwarz inequality for 2-positive maps, we have that

$$\begin{aligned}\|\mu_k(Q^*(X)Q(X))\|\|\mu_k(1)\| &\geq \|\mu_k(Q(X))\mu_k(Q^*(X))\|\\ &= \|\mu_k(Q(X))\|^2 > M^{2n}\end{aligned}$$

Since $\mu_k(1) = 1$, we may assume that our monomial $P(X) = Xb_1Xb_2\cdots b_{n-1}X^2b_{n-1}^*X\cdots b_1^*X$ has the property that $\mu_k(P(X)) > M^{2n}$. Define an element $B \in M_{2n}(\mathcal{B})$ by

$$B = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & \cdots & 0\\ 1 & 0 & b_1 & 0 & 0 & 0 & 0 & \cdots & 0\\ 0 & b_1^* & 0 & 1 & 0 & 0 & 0 & \cdots & 0\\ 0 & 0 & 1 & 0 & b_2 & 0 & 0 & \cdots & 0\\ 0 & 0 & 0 & b_2^* & 0 & 1 & 0 & \cdots & 0\\ 0 & 0 & 0 & 0 & 1 & 0 & b_3 & \cdots & 0\\ & \vdots & & & \vdots & & & & \vdots\\ 0 & 0 & 0 & 0 & 0 & \cdots & b_{n-1}^* & 0 & 1\\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 \end{pmatrix}.$$

That is, the superdiagonal alternates between 1 and b_i, the subdiagonal alternates between 1 and b_i^*. Now, let $0 < \epsilon, \delta$ and

$$B_{\delta,\epsilon} = \delta B + \epsilon\left(\sum_{i=1}^{2n-1} e_{i,i}\right) + \frac{e_{2n,2n}}{\delta^{n-1}}$$

where ϵ is arbitrarily small and δ is chosen so that $B_{\delta,\epsilon}$ is a strictly positive element. Moreover, we have that

(5)
$$\begin{aligned} e_{1,1}(B_{\delta,\epsilon}(X\otimes 1_{2n})B_{\delta,\epsilon})^{2n}e_{1,1} &= e_{1,1}B_{\delta,\epsilon}[(X\otimes 1_{2n})B_{\delta,\epsilon]}^2]^{2n-1}(X\otimes 1_{2n})B_{\delta,\epsilon}e_{1,1} \\ &= P(X)+O(\max(\delta,\epsilon)). \end{aligned}$$

To see this, note that a non-trivial contribution to (5) must be of the form

$$b_{1,2}Xb_{2,j_3}b_{j_3,j_4}Xb_{j_4,j_5}X\cdots b_{j_{4n-2},j_{4n-1}}b_{j_{4n-1},2}Xb_{2,1}$$

where $b_{i,j}$ denotes the i,j entry of $B_{\delta,\epsilon}$. Now, such a non-zero term is *not* $O(\max(\delta,\epsilon))$ means that $b_{j_\ell,j_{\ell+1}}$ must equal $b_{2n,2n}$ for two distinct ℓ. However, the only possible way for this to occur is if $j_k=k$ for $k=2,\ldots,2n$, $j_{2n}=j_{2n+1}=j_{2n+2}=2n$ and $j_p=4n+2-p$ for $p=2n+2,\ldots,4n-1$.
By assumption, there exists a state $\phi\in\mathcal{B}^*$ such that $\phi(\mu_k(P(X)))>M^{2k}$. Thus, for ϵ small enough, we have that

$$\phi_{1,1}\circ(\mu_k\otimes 1_{2n})((B_{\delta,\epsilon}(X\otimes 1_{2n})B_{\delta,\epsilon})^{2n})>M^{2n} \tag{6}$$

(here $\phi\otimes e_{1,1}=\phi_{1,1}$). This implies that the scalar valued Cauchy transform associated to this random variable,

$$G^{\delta,\epsilon}_{\mu_k}(z)=\phi_{1,1}\circ(\mu_k\otimes 1_{2n})((z1_{2n}-B_{\delta,\epsilon}(X\otimes 1_{2n})B_{\delta,\epsilon})^{-1})$$

arises from a measure whose support has non-trivial intersection with $\mathbb{R}\setminus[-M,M]$, whereas the (similarly defined) $G^{\delta,\epsilon}_{\mu}$ has support contained in $[-M,M]$ (since its moments have growth rate smaller than powers of M). Using Stieltjes inversion, this implies that

$$\lim_{t\downarrow 0}-\Im G^{\delta,\epsilon}_{\mu_k}(x+it)>0 \tag{7}$$

for some $x>M$ (or the limit simply does not exist in the atomic case).
Calculating the imaginary part of this Cauchy transform, we have

$$\begin{aligned} \Im([\mu_k((z1_{2n}-B_{\delta,\epsilon}XB_{\delta,\epsilon})^{-1})]^{-1}) &= B^{-1}_{\delta,\epsilon}\Im([\mu_k(B^{-2}_{\delta,\epsilon}z-X)^{-1}]^{-1})B^{-1}_{\delta,\epsilon} \\ &= B^{-1}_{\delta,\epsilon}\Im F^{(n)}_{\mu_k}(zB^{-2}_{\delta,\epsilon})B^{-1}_{\delta,\epsilon} \\ &\le B^{-1}_{\delta,\epsilon}\Im F^{(n)}_{\mu}(zB^{-2}_{\delta,\epsilon})B^{-1}_{\delta,\epsilon} \\ &= \Im([\mu((z1_{2n}-B_{\delta,\epsilon}XB_{\delta,\epsilon})^{-1})]^{-1}) \end{aligned} \tag{8}$$

where the inequality follows from the fact that $F_\mu=F^{\circ k-1}_{\mu_k}\circ F_{\mu_k}$ and F-transforms increase the imaginary part.
Rewriting the right hand side of (8), we have that

$$\begin{aligned} &\Im([\mu((z1_{2n}-B_{\delta,\epsilon}XB_{\delta,\epsilon})^{-1})]^{-1}) \\ &\quad=[\mu((z1_{2n}-B_{\delta,\epsilon}XB_{\delta,\epsilon})^{-1})^*]^{-1}\cdot \\ &\qquad\qquad \Im(\mu((z1_{2n}-B_{\delta,\epsilon}XB_{\delta,\epsilon})^{-1}))[\mu((z1_{2n}-B_{\delta,\epsilon}XB_{\delta,\epsilon})^{-1})]^{-1} \\ &\quad=F^{\delta,\epsilon}_{\mu}(z)^*\Im(F^{\delta,\epsilon}_{\mu}(z))F^{\delta,\epsilon}_{\mu}(z) \end{aligned} \tag{9}$$

We conclude that

$$\Im([\mu((z1_{2n} - B_{\delta,\epsilon} X B_{\delta,\epsilon})^{-1})]^{-1}) \leq F_\mu^{\delta,\epsilon}(z)^* \Im(F_\mu^{\delta,\epsilon}(z)) F_\mu^{\delta,\epsilon}(z). \tag{10}$$

Since $F_\mu^{\delta,\epsilon}$ extends to $\mathbb{R} \setminus [-M, M]$

$$\lim_{t \downarrow 0} G_\mu^{\delta,\epsilon}(x + it)$$

converges to a positive element in $\mathcal{B}$ and

$$\lim_{t \downarrow 0} \Im(F_\mu^{\delta,\epsilon}(x + it)) \to 0$$

it follows that the right hand side or (10) converges to 0 in norm, contradicting (7). This completes our proof. □

PROPOSITION 3.2. Let μ, μ_k be as in the preceding proposition. We have that $F_{\mu_k} \to Id$ in norm as $k \uparrow \infty$ uniformly on $M_n^{+,\epsilon}(\mathcal{B})$, and this convergence is also uniform over n . Moreover, the functions $F_{\mu_k}^{(n)}(b^{-1}) - b^{-1}$ and $F_{\mu_k}^{(n)}(b^{-1})^{-1}$ extend analytically to $B_r^{nc}(0)$, where the radius r is dependent only on M from Proposition 3.1, and satisfy

$$F_{\mu_k}^{(n)}(b^{-1}) - b^{-1} \to 0_n \tag{11}$$

$$F_{\mu_k}^{(n)}(b^{-1})^{-1} = H_{\mu_k}^{(n)}(b) \to b \tag{12}$$

where this convergence is uniform on $B_r^{nc}(0)$.

Proof. Consider the Nevanlinna representations of each of these functions

$$F_{\mu_k}^{(n)}(b) = \alpha_k \otimes 1_n + b - G_{\rho_k}^{(n)}(b) \tag{13}$$

defined in Theorem 2.1 , where we have adopted the notation that $\mu = \mu_1$. We claim that the distributions ρ_k share a common exponential bound N for all $k \in \mathbb{N}$.
To prove this claim, first observe that, by Theorem 4.1 in [Wil13], there exist distributions ν_k such that

$$b - F_{\mu_k}^{(n)}(b) = \varphi_{\nu_k}^{(n)}(b) = -\alpha_k \otimes 1_n + G_{\rho_k}^{(n)}(b).$$

Moreover, it was shown in [PV13] that if the ν and the ν_k have a common exponential bound N then the distributions ρ and ρ_k have a common exponential bound $N^2 + 1$. Focusing on the ν_k, we may manipulate equations 13 to conclude that

$$\mathcal{R}_{\nu_k}(b^{-1}) = \varphi_{\nu_k}(b) = b^{-1} - F_{\mu_k}(b^{-1}). \tag{14}$$

Now, expand the moment series

$$F_{\mu_k}^{(n)}(b^{-1})^{-1} = H_{\mu_k}^{(n)}(b) = \sum_{p=0}^{\infty} \mu_k((bX)^p b). \tag{15}$$

Note that Proposition 3.1 implies that this function is convergent and uniformly bounded for $b \in B_r^{nc}(0)$, independent of k.
Observe that the moment generating function satisfies

$$[H_{\mu_k}^{(n)}(b)]^{-1} = b^{-1} - \mu_k(X) + \mu_k(X) b \mu_k(X) - \mu_k(XbX) + \cdots = b^{-1} + f^{(n)}(b, X) \tag{16}$$

where $f^{(n)}(b,X)$ is analytic in b and converges for $\|b\|$ small, where the radius of convergence is only dependent on M. Thus, $[H^{(n)}_{\mu_k}(b)]^{-1} - b^{-1}$ extends to a neighborhood of 0 whose radius is independent of n and k and agrees with $F^{(n)}_{\mu_k}(b^{-1}) - b^{-1}$ when b is invertible. Moreover, these observations, combined with (14) imply that the functions $\mathcal{R}_{\nu_k}$ have a common $R, C > 0$ such that the functions extend to a common domain $B^{nc}_R(0)$ with a common bound C. Now a careful look at the Kantorovich argument in part II of the proof of Theorem 4.1 in [Wil13] allows us to conclude that the exponential bound on the distributions ν_k depend only on R, proving our claim. Recall that $F_{\mu_k} \circ \cdots \circ F_{\mu_k} = F_\mu$ we have that

$$G^{(n)}_\rho(b) = G^{(n)}_{\rho_k}(b) + G^{(n)}_{\rho_k} \circ F^{(n)}_{\mu_k}(b) + \cdots + G^{(n)}_{\rho_k} \circ \underbrace{F^{(n)}_{\mu_k} \circ \cdots \circ F^{(n)}_{\mu_k}}_{k-1 \ times}(b) \tag{17}$$

Letting $b = z1_n$ for $z \in \mathbb{C}$, we have that

$$\begin{aligned}
\lim_{|z|\uparrow\infty} zH^{(n)}_\rho\left(\frac{1}{z}1_n\right) &= \lim_{|z|\uparrow\infty} zG^{(n)}_\rho(z1_n) \\
&= \lim_{|z|\uparrow\infty} \sum_{\ell=1}^{k-1} zG^{(n)}_{\rho_k} \circ (F^{(n)}_{\mu_k})^{\circ\ell}(z1_n) \\
&= \lim_{|z|\uparrow\infty} \sum_{\ell=1}^{k-1} zH^{(n)}_{\rho_k}\left([(F^{(n)}_{\mu_k})^{\circ\ell}(z1_n)]^{-1}\right) \\
&= \lim_{|z|\uparrow\infty} \sum_{\ell=1}^{k-1} zH^{(n)}_{\rho_k} \circ G^{(n)}_{\nu_\ell}(z1_n) \\
&= \lim_{|w|\downarrow 0} \sum_{\ell=1}^{k-1} \frac{1}{w} H^{(n)}_{\rho_k} \circ G^{(n)}_{\nu_\ell}\left(\frac{1}{w}1_n\right) \\
&= \lim_{|w|\downarrow 0} \sum_{\ell=1}^{k-1} \frac{1}{w} H^{(n)}_{\rho_k} \circ H^{(n)}_{\nu_\ell}(w1_n)
\end{aligned}$$

where $[(F^{(n)}_{\mu_k})^{\circ\ell}]^{-1} = G_{\nu_\ell}$ is the Cauchy transform of a distribution $\nu_\ell \in \Sigma_0$ (this follows from Theorem 2.1). Moreover, we have that

$$\lim_{|w|\downarrow 0} \frac{1}{w} H^{(n)}_{\nu_\ell}(w1_n) = 1_n$$

so that, passing to limits and utilizing the chain rule and the fact that $H^{(n)}_{\nu_\ell}(0_n) = 0_n$, we have that

$$\delta H^{(n)}_\rho(0_n; 1_n) = k\delta H^{(n)}_{\rho_k}(0_n; 1_n)$$

Utilizing the main result in our appendix, Theorem A.1, we conclude that

$$\rho(1) = \mu(X^2) = k\mu_k(X^2) = k\rho_k(1). \tag{18}$$

so that $\rho_k(1) = O(1/k)$.

Now, assume that $b \in M_n^{+,\epsilon}(\mathcal{B})$. We claim that $\|b^{-1}\| \leq 1/\epsilon$. Indeed, observe that, for $b = x + iy$ with $y > \epsilon 1_n$,

$$b = \sqrt{y}(i + (\sqrt{y})^{-1}x(\sqrt{y})^{-1})\sqrt{y} \tag{19}$$

(it follows easily from this equation that b is invertible, but this is known). Thus,

$$b^{-1} = (\sqrt{y})^{-1}(i + (\sqrt{y})^{-1}x(\sqrt{y})^{-1})^{-1}(\sqrt{y})^{-1}. \tag{20}$$

Now, utilizing the spectral mapping theorem and the fact that the spectral radius agrees with the norm for normal operators, we have that $\|(\sqrt{y})^{-1}\| \leq (\sqrt{\epsilon})^{-1}$. Moreover, since $i + (\sqrt{y})^{-1}x(\sqrt{y})^{-1}$ is normal and has spectrum with imaginary part larger than 1, we have that $(i + (\sqrt{y})^{-1}x(\sqrt{y})^{-1})^{-1}$ is normal and, by the same spectral considerations, has norm bounded by 1. These observations, combined with (19) imply our claim.

Thus, for $b \in M_n^{+}(\mathcal{B})$, we have

$$\begin{aligned}
\|F_{\mu_k}^{(n)}(b) - b\| &\leq \|\alpha_k\| + \|(\rho_k \otimes 1_n)((b - X)^{-1}\| \\
&\leq \|\alpha\|/k + \|(b - X)^{-1}\|\|(\rho_k \otimes 1_n)(1_n)\| \\
&\leq \frac{\|\alpha\|}{k} + \frac{\|\rho_k(1)\|}{\epsilon} = \frac{\|\alpha\| + \rho(1)/\epsilon}{k}
\end{aligned}$$

and the right hand side converges to zero uniformly over $M_n^{+,\epsilon}(\mathcal{B})$, independent of n. Regarding the second part of our Proposition, we first observe that each of the moments of μ_k converges to 0. Indeed, utilizing the Schwarz inequality for 2-positive maps as well as Proposition 3.1, we have that

$$\begin{aligned}
\|\mu_k(Xb_1Xb_2X \cdots b_\ell X)\|^2 &\leq \|\mu_k(X^2)\|\|\mu_k(Xb_\ell^* X \cdots b_2^* X b_1^* b_1 X b_2 X \cdots b_\ell X)\| \\
&\leq \frac{\|\mu(X^2)\| M^{2\ell}\|b_1\|^2\|b_2\|^2 \cdots \|b_\ell\|^2}{k}
\end{aligned}$$

Moreover, the tail of the series expansion of $f^{(n)}(b, X)$ is bounded in norm independent of n and k . the individual entries all go to 0 so the we conclude that $f^{(n)}(b, X) \to 0$ uniformly on $b \in B_r^{nc}(0)$ as $k \uparrow \infty$ so that we can immediately conclude that (12) holds. This completes our proof. □

We next prove a differentiation result for vector valued functions. We adapt a proof found in [BPo78] of a similar result for complex functions.

Proposition 3.3. Let $\mathcal{A}$ and $\mathcal{B}$ denote unital Banach algebras. Consider an open subset $\Omega \subset \mathcal{A}$. Let $f_t : \Omega \mapsto \mathcal{B}$ for all $t \geq 0$ be a composition semigroup of analytic functions. Assume that for every $b' \in \Omega$, there exists a $\delta > 0$ such that

(a) $\lim_{t \downarrow 0} f_t(b) - b \to 0$ uniformly over $b \in B_\delta(b')$
(b) For any $T > 0$, we have that $f_t(b) - b$ is uniformly bounded over $b \in B_\delta(b')$ and $t \in [0, T]$.

Then, there exists an analytic $\Phi : \Omega \mapsto \mathcal{B}$ such that

$$\frac{df_t(b)}{dt} = -\Phi(f_t(b)). \tag{21}$$

Proof. Fix $b' \in \Omega$. We first claim that there exists an $\alpha > 0$ such that

$$\|f_{2t}(b) - 2f_t(b) + b\| \leq \frac{1}{10}\|f_t(b) - b\|. \tag{22}$$

for all $t \in [0, \alpha]$ and $b \in B_{\delta/2}(b')$ where the value of δ comes from the statement . Indeed, fix $b \in B_{\delta/2}(b')$. We first consider the simple case when there exists a sequence $t_n \downarrow 0$ such that $f_{t_n}(b) = b$. Since $\{f_t\}$ form a composition semigroup, this property then holds for a dense set of t's, and by continuity assumption in part (a), for all $t > 0$. So (22) holds trivially.

Thus, suppose that $f_t(b) \neq b$ for $t \in [0, \alpha]$. Define a family of complex functions g_t through the following equalities:

$$h_t := \frac{f_t(b) - b}{\|f_t(b) - b\|};\ g_t(\zeta) := f_t(b + \zeta h_t) - b : B_{\delta/2}(0) \mapsto \mathcal{B}.$$

where $B_{\delta/2}(0)$ refers to the neighborhood of zero in the complex plane. Note that, since we are taking a ball of radius $\delta/2$, we may define h_t for all such b provided that our choice of α is small enough.

Consider the vector valued complex integral

$$\int_0^{\|f_t(b)-b\|} \frac{d}{d\zeta}[g_t(\zeta) - \zeta h_t] d\zeta. \tag{23}$$

By (a) and the Cauchy estimates in Theorem 2.2, the integrand can be made arbitrarily small for t small. By the fundamental theorem, this integral is equal to

$$\begin{aligned} &g_t(\|f_t(b) - b\|) - g_t(0) - (f_t(b) - b) = \\ &\qquad = f_t(b + (f_t(b) - b)) - b - 2(f_t(b) - b) = f_{2t}(b) - 2f_t(b) + b. \end{aligned}$$

Using our bound on the integrand, equation (22) follows immediately.

We now use (22) to prove that for $\alpha > 0$ there exists an $M > 0$ such that

$$\|f_t(b) - b\| \leq Mt^{2/3} \tag{24}$$

for all $t \in [0, \alpha]$ and $b \in B_{\delta/2}(b')$. Indeed, pick $t \in [0, \alpha]$ and $m \in \mathbb{N}$ such that $2^m t \leq \alpha < 2^{m+1}t$. Note that inequality (22) and the triangle inequality imply that

$$2\|f_t(b) - b\| - \|f_{2t}(b) - b\| \leq \|f_{2t}(b) - 2f_t(b) + b\| \leq \frac{1}{10}\|f_t(b) - b\|$$

so that

$$\|f_t(b) - b\| \leq \frac{10}{19}\|f_{2t}(b) - b\| \leq 2^{-2/3}\|f_{2t}(b) - b\| \tag{25}$$

Using this estimate inductively, we have

$$\|f_t(b) - b\| \leq 2^{-2/3}\|f_{2t}(b) - b\| \leq \cdots \leq 2^{-2m/3}\|f_{2^m t}(b) - b\| = t^{2/3}\left(\frac{1}{2^m t}\right)^{2/3} M'$$

where M' is a bound on $\|f_s(b) - b\|$ for $s \leq 2$ which exists by (b). Equation (24) follows with $M = 2^{2/3}M'/\alpha$.

Now, revisiting the argument for (22), inequality (24) implies that the integrand in (23) has bound equal to

$$2Mt^{2/3}$$

as a result of the Cauchy estimates. Thus, we have the following:

$$\|f_{2t}(b) - 2f_t(b) + b\| \leq 2t^{2/3}\|f_t(b) - b\| \leq 2Mt^{4/3}. \tag{26}$$

We may further conclude that

$$\left\| \frac{f_{2t}(b) - b}{2t} - \frac{f_t(b) - b}{t} \right\| \leq Mt^{1/3} \tag{27}$$

Thus, we have that

$$\lim_{k \uparrow \infty} 2^k (f_{2^{-k}}(b) - b) \tag{28}$$

converges uniformly on $B_{\delta/2}(b')$ and we refer to this limit as $-\Phi(b)$.
Using (27), we note that Φ is locally bounded. Indeed, we have that

$$\begin{aligned}\|2^p(f_{1/2^p}(b) - b) + \Phi(b)\| &\leq \sum_{k=p}^{\infty} \|2^k(f_{1/2^k}(b) - b) - 2^{k+1}(f_{1/2^{k+1}}(b) - b)\| \\ &\leq \frac{M}{2} \sum_{k=p}^{\infty} \left(\frac{1}{2^{1/3}}\right)^k = MC(p). \qquad (29)\end{aligned}$$

for all $b \in B_{\delta/2}(b')$. Local boundedness of Φ follows since $(f_{1/2^p}(b) - b)$ is locally bounded. Also note that $C(p) \to 0$ as $p \uparrow \infty$.
Regarding analyticity of Φ, consider a state $\varphi \in \mathcal{B}^*$, $b \in B_{\delta/2}(b')$, and an element $h \in \mathcal{B}$ with $\|h\| \leq 1$. We define complex maps

$$H_m(z) : B_{\delta/2}(0) \subset \mathbb{C} \to \mathbb{C}$$

for $m \geq 0$ through the equalities:

$$H_0(z) := \varphi \circ \Phi(b + zh); \;\; H_m(z) := 2^m \varphi \circ (f_{2^{-m}}(b + zh) - (b + zh)).$$

By (28), $H_m \to H_0$ for $z \in B_{\delta/2}(0)$, and by (29), the limit is bounded on this set. Thus, H_0 is analytic in z. By Dunford's theorem ([Dun38]), it follows that $\Phi(b + zh)$ is analytic in z and, therefore, Gâteaux differentiable. As this function is locally bounded, it is analytic.
Regarding (21), observe that $\{f_t(b)\}_{t \geq 0}$ is compact since it is the continuous image of $[0, t]$. As (a) and (b) hold on neighborhoods of every point in this set, taking a finite cover, we have that (a) and (b) holds uniformly on a neighborhood of this set and, after a close look at the relevant constants, (29) is also maintained on this set. Now, fix $t \geq 0$ and let $\ell_p/2^p \to t$ as $p \uparrow \infty$.

$$\begin{aligned} f_t(b) - b &= (f_t(b) - f_{t-\ell_p/2^p}(b)) + \sum_{j=1}^{\ell_p} (f_{j/2^p}(b) - f_{(j-1)/2^p}(b)) \\ &= (f_t(b) - f_{t-\ell_p/2^p}(b)) + \sum_{j=1}^{\ell_p} \frac{1}{2^p} (2^p [f_{j/2^p}(b) - f_{(j-1)/2^p}(b)]) \end{aligned}$$

As $p \uparrow \infty$,

$$f_t(b) - f_{t-\ell_p/2^p}(b) = f_{\ell_p/2^p} \circ f_{t-\ell_p/2^p}(b) - f_{t-\ell_p/2^p}(b) \to 0$$

since (a) holds on the entire path. Moreover, the remaining summand is simply a Riemann sum approximation of a sequence of functions converging uniformly to $-\Phi\circ f_s(b)$ for $s \in [0,t]$. The following equation follows immediately:

$$f_t(b) = b - \int_0^t \Phi \circ f_s(b) ds.$$

We conclude that (21) holds, completing our proof. □

Corollary 3.4. Let $\mathcal{A}$ and $\mathcal{B}$ denote Banach algebras and $\Omega \subset \sqcup_{n=1}^{\infty} M_n(\mathcal{A})$ a non-commutative set. Let $F_t : \Omega \mapsto \sqcup_{n=1}^{\infty} M_n(\mathcal{B})$ for all $t \geq 0$ and assume that they form a composition semigroup of analytic non-commutative functions. Assume that, for each n, the composition semigroup of vector valued analytic functions $\{F_t^{(n)}\}_{t\geq 0}$ satisfies the hypotheses of Proposition 3.3. Then there exists an analytic, noncommutative map $\Phi : \Omega \mapsto \sqcup_{n=1}^{\infty} M_n(\mathcal{B})$ such that

$$\frac{dF_t^{(n)}(b)}{dt} = -\Phi^{(n)}(F_t^{(n)}(b)) \tag{30}$$

for all $n \in \mathbb{N}$, $b \in \Omega_n$.
Moreover, if we strengthen these assumptions so that, for any n and $b \in M_n(\mathcal{B})$, there exists a $\delta > 0$ with

(a) $\lim_{t\downarrow 0} F_t - Id \to 0$ uniformly over $B_\delta^{nc}(b)$.
(b) For any $T > 0$, we have that $f_t(b) - b$ is uniformly bounded on $B_\delta^{nc}(b)$ and $t \in [0,T]$.

then Φ is uniformly analytic.

Proof. We showed in Proposition 3.3 this map Φ exists. We must show that it is a non-commutative function. However, this is immediate since, for $b_1 \in M_n(\mathcal{B})$ and $b_2 \in M_p(\mathcal{B})$, we have

$$\begin{aligned}\Phi^{(n+p)}(b_1 \oplus b_2) &= \lim_{k\uparrow\infty} 2^k (F_{2^{-k}}^{(n+p)}(b_1 \oplus b_2) - b_1 \oplus b_2) \\ &= \lim_{k\uparrow\infty} 2^k ([F_{2^{-k}}^{(n)}(b_1) - b_1] \oplus [F_{2^{-k}}^{(n)}(b_2) - b_2]) \\ &= \Phi^{(n)}(b_1) \oplus \Phi^{(p)}(b_2).\end{aligned}$$

A similar proof shows that it also satisfies the defining invariance property so that our first claim holds.
With respect to the uniform analyticity, we refer to the proof of Proposition 3.3. Observe that inequality (22) holds for α small enough. This α is only dependent on the convergence of the integrand in (23). This converges to 0 uniformly on $B_\delta^{nc}(b)$ by assumption (a) and the same Cauchy estimate so that the choice of α is also uniform on this set. Moreover, the constant M in (24) is equal to $2^{2/3}M'/\alpha$ where M' is the upper bound on $F_s - Id$ for $s \leq \alpha$. Assumption (b) implies that this bound is uniform on $B_\delta^{nc}(b)$. Thus, inequality (29) holds on all of this set, implying uniform analyticity. □

Theorem 3.5. Let $\{F_t\}_{t\in\mathbb{Q}^+}$ denote a composition semigroup of non-commutative functions $F_t : H^+(\mathcal{B}) \mapsto H^+(\mathcal{B})$ such that

(i) $\|F_t^{(n)}(b) - b\| \to 0$ uniformly on $M_n^{+,\epsilon}(\mathcal{B})$ for all $\epsilon > 0$, independent of n as $t \downarrow 0$.
(ii) For any $\alpha, \epsilon > 0$ and sequence $b_k \in \Gamma_{\alpha,\epsilon}^{(n)}$ with $\|b_k^{-1}\| \downarrow 0$, we have that $b_k^{-1} F_t^{(n)}(b_k) \to 1_n$ as $k \uparrow \infty$
(iii) $\Im F_t^{(n)}(b) \geq \Im b$ for all $b \in M_n^+(\mathcal{B})$ and $t \geq 0$.

Then $\{F_t\}_{t \in \mathbb{Q}^+}$ extends to a semigroup $\{F_t\}_{t \geq 0}$ and the map Φ from Proposition 3.4 is an element of $\tilde{\Lambda}$.
Since, by Proposition 3.2, the conditions above are satisfied by F-transforms, this implies that a $\rhd$-infinitely divisible distribution μ as in Definition 2.4 can be realized as $\mu = \mu_1$ for a monotone convolution semigroup $\{\mu_t\}_{t \geq 0}$. For such a semigroup, $\Phi \in \Lambda$.
Conversely, given a map $\Phi \in \tilde{\Lambda}$ we may construct a semigroup of non-commutative functions satisfying the hypotheses above as well as the differential equation

$$\frac{dF_t(b)}{dt} = -\Phi(F_t(b)) \tag{31}$$

If $\Phi \in \Lambda$ then the semigroup arises from a $\rhd$-infinitely divisible distribution.

We shall refer to this element Φ as the *generator* or the semigroup $\{F_t\}_{t \geq 0}$.

Proof. First, let $\Phi \in \tilde{\Lambda}$. We will produce the semigroup it generates by the method of successive approximations.
Consider a sequence of non-commutative functions $\{f_k(t, \cdot)\}_{t \geq 0,\ k \in \mathbb{N}}$ defined as follows:

$$f_1^{(n)}(t, b) = b; \;\; f_{k+1}^{(n)}(t, b) = b - \int_0^t \Phi(f_k^{(n)}(s, b))ds. \tag{32}$$

We claim that $f_k(t, \cdot)$ is convergent and satisfies the semigroup property with generator Φ.
Observe that since Φ is uniformly bounded by a constant M on set $M_n^{+,\epsilon/2}(\mathcal{B})$ and $f_k(t, \cdot)$ maps the set $M_n^{+,\epsilon}(\mathcal{B})$ to itself since

$$\Phi : H^+(\mathcal{B}) \mapsto H^-(\mathcal{B})$$

we have that

$$\Im f_k^{(n)}(t, b) \geq \Im(b). \tag{33}$$

By (3), this implies that $f_k^{(n)}(t, \cdot)$ is Lipschitz on the set $B_{\epsilon/2}(b) \subset M_n^{+,\epsilon/2}(\mathcal{B})$ for all $b \in M_n^{+,\epsilon}(\mathcal{B})$, and the Lipschitz constant L is uniform over both k, b and bounded t. Moreover, we may extend the Lipschitz inequality

$$\|f_k(t, b) - f_k(t, b')\| \leq L\|b - b'\|$$

to all $b, b' \in M_n^{+,\epsilon}(\mathcal{B})$ by taking a path $b + s(b' - b)$ for $s \in [0, 1]$ and using the Lipschitz estimate on intervals of distance $\epsilon/2$ since the distances are additive on this path. Using this Lipschitz estimate in the integrand of (32), we conclude that

$$\|f_2^{(n)}(t, b) - f_1^{(n)}(t, b)\| = t\|\Phi(b)\| \leq tML \tag{34}$$

and we may conclude that

$$\begin{aligned}\|f_3^{(n)}(t,b) - f_2^{(n)}(t,b)\| &= \left\|\int_0^t [\Phi(f_2^{(n)}(s,b)) - \Phi(f_1^{(n)}(s,b))]ds\right\| \\ &\leq L\left\|\int_0^t [f_2^{(n)}(s,b) - f_1^{(n)}(s,b)]ds\right\| \\ &\leq L\int_0^t [LMs]ds \leq \frac{t^2L^2M}{2}\end{aligned}$$

Continuing inductively, we have that

$$\|f_{k+1}^{(n)}(t,b) - f_k^{(n)}(t,b)\| \leq \frac{M(Lt)^{k+1}}{L(k+1)!}. \tag{35}$$

For any choice of $t \in [0,\alpha]$, we have that

$$f_{N+1}^{(n)}(t,b) - b = \sum_{k=0}^{N}\left(f_{k+1}^{(n)}(t,b) - f_k^{(n)}(t,b)\right) \tag{36}$$

is a convergent series as $N \uparrow \infty$ and we may conclude that $f_N(t,\cdot)$ converges to a function $f(t,\cdot)$ uniformly on $M_n^{+,\epsilon}(\mathcal{B})$, independent of n.
It is clear that $f(t,\cdot)$ satisfies (31). Regarding the asymptotics, let $\alpha, \epsilon > 0$ and fix a sequence $b_\ell \in \Gamma_{\alpha,\epsilon}^{(n)}$ with $\|b_\ell^{-1}\| \downarrow 0$. Note that $b_\ell^{-1}f_1^{(n)}(t,b_\ell) \equiv 1_n$ and satisfies $\|f_1^{(n)}(t,b_\ell)\|^{-1} \downarrow 0$ as $\|b_\ell^{-1}\| \downarrow 0$. We claim $b_\ell^{-1}f_k^{(n)}(t,b_\ell) \to 1_n$ and satisfies $\|f_k^{(n)}(t,b_\ell)\|^{-1} \downarrow 0$ as $\|b_\ell^{-1}\| \downarrow 0$ for all k, uniformly over $t \in [0,\alpha]$.
Proceeding by induction, we have that for fixed k

$$b_\ell^{-1}f_{k+1}^{(n)}(t,b_\ell) = 1_n - \int_0^t [b_\ell^{-1}f_k^{(n)}(s,b_\ell)](f_k^{(n)}(s,b_\ell))^{-1}\Phi(f_k^{(n)}(s,b_\ell))ds. \tag{37}$$

We bound the integrand by

$$\|[b_\ell^{-1}f_k^{(n)}(s,b_\ell)]\|\|(f_k^{(n)}(s,b_\ell))^{-1}\Phi(f_k^{(n)}(s,b_\ell))\|$$

which converges to 0 uniformly over $s \in [0,\alpha]$ by induction, so that (37) converges to 1_n. Moreover,

$$\|[f_{k+1}^{(n)}(t,b_\ell)]^{-1}\| \leq \|b_\ell^{-1}\|\|b_\ell[f_{k+1}^{(n)}(t,b_\ell)]^{-1}\| \to 0.$$

Thus, each $f_k(t,\cdot)$ has the appropriate asymptotics and, since $f(t,\cdot)$ is a uniform limit of these functions on $M_n^{+,\epsilon}$, our claim holds Condition (iii) follows from (33).
In order to complete our proof, we further assume that $\Phi \in \Lambda$ and prove that the functions $f(t,\cdot)$ are in fact the F-transforms of noncommutative distributions $\mu_t \in \Sigma_0$. To do so we must show that the function $f(t,b^{-1})^{-1}$ has a uniformly analytic extension to a neighborhood of 0 for all $t \geq 0$. Note that, since $\Phi \in \Lambda$, there exists a $\delta > 0$ and constants $M, L > 0$ such that $\Phi^{(n)}(b^{-1})$ extends to $B_\delta^{nc}(0)$ with upper bound M and Lipschitz constant L.
Now fix $\alpha > 0$. We claim that, for $\gamma > 0$ small enough we have that $f_k^{(n)}(t,b^{-1})^{-1}$ extends to $B_\gamma(0_n) \subset M_n(\mathcal{B})$ for all n and satisfies $f_k^{(n)}(t,b^{-1})^{-1} \in B_\delta(0_n)$ for all

$b \in B_\gamma(0_n)$. Choose any $t \in [0, \alpha]$ and $b \in B_\gamma(0_n)$ where $\gamma < \delta$ is yet unspecified. We have

$$\begin{aligned}\|f_2^{(n)}(t, b^{-1})^{-1} - f_1^{(n)}(t, b^{-1})^{-1}\| &= \left\| \left[\left(1_n - \int_0^t b\Phi(b^{-1}) ds \right)^{-1} - 1_n \right] b \right\| \\ &\leq \sum_{n=1}^{\infty} \left\| \int_0^t b\Phi(b^{-1}) ds \right\|^n \|b\| \\ &\leq \gamma \sum_{n=1}^{\infty} (\gamma M \alpha)^n \\ &= \frac{\gamma^2 M \alpha}{1 - \gamma M \alpha}\end{aligned}$$

Deriving a similar inequality for general k, we have that

(38)

$$\begin{aligned}&\|f_{k+1}^{(n)}(t, b^{-1})^{-1} - f_k^{(n)}(t, b^{-1})^{-1}\| \\ &= \left\| \left(b^{-1} - \int_0^t \Phi \circ f_k^{(n)}(s, b^{-1}) ds \right)^{-1} - \left(b^{-1} - \int_0^t \Phi \circ f_{k-1}^{(n)}(s, b^{-1}) ds \right)^{-1} \right\| \\ &= \left\| \left(1_n - \int_0^t b\Phi(f_k^{(n)}(t, b^{-1})) \right)^{-1} \left(b \int_0^t \Phi(f_{k-1}^{(n)}(t, b^{-1})) - \Phi(f_k^{(n)}(t, b^{-1})) \right) \right. \\ &\qquad \left. \left(1_n - \int_0^t b\Phi(f_{k-1}^{(n)}(t, b^{-1})) \right)^{-1} b \right\| \\ &\leq \left(\frac{1}{1 - \gamma M \alpha} \right)^2 (\gamma^2 L \alpha) \|f_k^{(n)}(t, b^{-1})^{-1} - f_{k-1}^{(n)}(t, b^{-1})^{-1}\|\end{aligned}$$

By induction, we have that

$$\|f_{k+1}^{(n)}(t, b^{-1})^{-1} - b\| = \sum_{\ell=1}^{k} \frac{M\gamma^{2\ell} L^{\ell-1} \alpha^\ell}{(1 - \gamma M \alpha)^{2\ell-1}}$$

This is convergent as $k \uparrow \infty$ for γ small and converges to 0 as $\gamma \downarrow 0$. Thus, for γ small enough, we have that $f_{k+1}^{(n)}(t, b^{-1}) \in B_\delta(0_n)$ for all k and n and, therefore, converges to a limit function on $B_\gamma(0_n)$ (since the differences in (38) are Cauchy). This limit function must agree with $f(t, \cdot)$ by analytic continuation. This completes our proof that $f(t, \cdot)$ is an F-transform for all t.

To address the converse, consider a semigroup $\{F_t\}_{t \in \mathbb{Q}^+}$ satisfying the (i) and (ii) in the statement of the theorem. First note that this easily extends to an $\mathbb{R}^+$ composition semigroup. Indeed, define $F_t(b) = \lim_{p/q \to t} F_{p/q}(b)$. To see that this is well defined, note that, as $p/q, p'/q' \to t$, we have

$$\|F_{p/q}^{(n)}(b) - F_{p'/q'}^{(n)}(b)\| = \|F_{p/q - p'/q'}^{(n)} \circ F_{p'/q'}^{(n)}(b) - F_{p'/q'}^{(n)}(b)\| \to 0$$

uniformly on $M_n^{+,\epsilon}(\mathcal{B})$ by property (i) and (iii) . It is immediate that this is a composition semigroup over $\mathbb{R}^+$ satisfying (i), (ii) and (iii).

By Corollary 3.4, this semigroup may be differentiated to produce a non-commutative function Φ . Regarding the asymptotics of Φ, consider the inequality

$$\|b^{-1}\Phi^{(n)}(b)\| \leq \left\| \frac{b^{-1}(F_t^{(n)}(b)-b)}{t} \right\| + \|b^{-1}\| \left\| \frac{(F_t^{(n)}(b)-b)}{t} - \Phi^{(n)}(b) \right\|. \tag{39}$$

Utilizing inequality (29) in the proof of Proposition 3.3 produces

$$\left\| \frac{(F_{2^N}^{(n)}(b)-b)}{2^N} - \Phi^{(n)}(b) \right\| \leq M \sum_{k=N+1}^{\infty} \left(\frac{1}{2^{1/3}} \right)^k \tag{40}$$

where this $M = 2M'/\alpha$. As was noted in the proof of Corollary 3.4, uniform convergence in the sense of (i) and (ii) implies a uniform bound on M. Thus, (40) converges to 0 uniformly on $M_n^{+,\epsilon}(\mathcal{B})$ so that, for fixed t small enough, second term on the right hand side of (39) is smaller than any $\delta > 0$ for $b \in M_n^{+,\epsilon}(\mathcal{B})$. Letting $b_k \in \Gamma_{\alpha,\epsilon}^{(n)}$ satisfy $\|b_k^{-1}\| \downarrow 0$, the first term on the right hand side of (39) converges to 0 by assumption (ii), and it follows that $\Phi \in \tilde{\Lambda}$.

If $\{F_t\}_{t\geq 0}$ arises from a $\rhd$-infinitely divisible measure, then it follows from Proposition 3.1 and Theorem 2.1 that $b_k^{-1}F_{\mu_t}^{(n)}(b_k) \to 1_n$ for any sequence $b_k \in M_n(\mathcal{B})$ with $\|b_k^{-1}\| \downarrow 0$ and a similar proof allows one to conclude that Φ satisfies condition (ii) in the definition of Λ.

It remains to show that Φ satisfies (i). However, Proposition 3.2 implies that there exists a fixed $r > 0$ such that each function $F_{\mu_t}^{(n)}(b^{-1}) - b^{-1}$ extends to $B_r(\{0\})$ and converges to 0 uniformly on this set. Thus, the strengthened hypotheses in Corollary 3.3 hold so that the non-commutative function defined by the equalities

$$\mathcal{R}^{(n)}(b) = \lim_{t\downarrow 0} \frac{F_{\mu_t}^{(n)}(b^{-1}) - b^{-1}}{t}$$

is uniformly analytic at 0 and, by continuation, is an extension of $\Phi^{(n)}(b^{-1})$ for each n. Thus, $\Phi \in \Lambda$, completing our proof. □

The following proposition establishes continuity in generating the semigroups, and may be useful in future applications.

Proposition 3.6. Assume that $\Phi_1, \Phi_2 \in \tilde{\Lambda}$ generate the semigroups of noncommutative functions $\{F_1(t,\cdot)\}_{t\geq 0}$ and $\{F_2(t,\cdot)\}_{t\geq 0}$. If we assume that $\|\Phi_1^{(n)}(b) - \Phi_2^{(n)}(b)\| < \epsilon$ for all $b \in B_\delta(b') \subset M_n(\mathcal{B})$, a ball of radius δ where $\Im(b') > \delta 1_n$, then we may conclude that $\|F_1^{(n)}(1,b) - F_2^{(n)}(1,b)\| < C\epsilon$ for all $b \in B_\delta(b')$ where C depends only on Φ_1.

Proof. To prove our claim, we first note that, by the vector-valued chain rule,

$$\frac{\delta^2 F_i^{(n)}(t,b)}{\delta t^2} = \delta\Phi^{(n)} \left(F_i^{(n)}(t,b), \frac{\delta}{\delta t} F^{(n)}(b,t) \right)$$

so that $F_i(t,b)$ is twice differentiable in t and has uniformly bounded derivative for $b \in H^{+,\epsilon}(\mathcal{B})$ and $t \in [0,1]$. We refer to the maximum of this bound over $i = 1,2$ as M_2.

Using the remainder estimates for the Taylor series associated to F_i, we have the following:

$$\|F_i(b, t+\gamma) - F_i(b,t) - \gamma\Phi(F_i(b,t))\| \leq \frac{M_2\gamma^2}{2} \tag{41}$$

Let $M_1 = \sup_{b \in M_n^{+,\epsilon}(\mathcal{B}),\, n \in \mathbb{N}} \|\delta\Phi^{(n)}(b,\cdot)\|$. Utilizing the estimate (41) with $\gamma = 1/N$, we produce the following inequalities:

$$\begin{aligned}
&\|F_1^{(n)}(b, t_0 + 1/N) - F_2^{(n)}(b, t_0 + 1/N)\| \\
&\leq \frac{M_2}{N^2} + \frac{1}{N}\|\Phi_1^{(n)}(F_1^{(n)}(b,t_0)) - \Phi_2^{(n)}(F_2^{(n)}(b,t_0))\| \\
&\quad + \|F_1^{(n)}(b,t_0) - F_2^{(n)}(b,t_0)\| \\
&\leq \frac{M_2}{N^2} + \frac{1}{N}\|\Phi_1^{(n)}(F_1^{(n)}(b,t_0)) - \Phi_1^{(n)}(F_2^{(n)}(b,t_0))\| \\
&\quad + \frac{1}{N}\|\Phi_1^{(n)}(F_2^{(n)}(b,t_0)) - \Phi_2^{(n)}(F_2^{(n)}(b,t_0))\| + \|F_1^{(n)}(b,t_0) - F_2^{(n)}(b,t_0)\| \\
&\leq \frac{M_2}{N^2} + \frac{\epsilon}{N} + \left(1 + \frac{M_1}{N}\right)\|F_1^{(n)}(b,t_0) - F_2^{(n)}(b,t_0)\|
\end{aligned}$$

Using this estimate inductively, we have that

$$\|F_1^{(n)}(b,1) - F_2^{(n)}(b,1)\| \leq \left(\frac{\epsilon}{N} + \frac{M_2}{N^2}\right)\sum_{k=0}^{N-1}\left(1 + \frac{M_1}{N}\right)^k \to \frac{e^{M_1} - 1}{M_1}\epsilon$$

where the convergence occurs as $N \uparrow \infty$. This implies our result.

□

4. The Bercovici-Pata Bijection.

DEFINITION 4.1. Let $(S, \prec)$ be a poset (partially ordered set). An *order* on S is an order-preserving bijection

$$f : (S, \prec) \to (\{1, 2, \ldots, |S|\}, <).$$

Denote by $o(S)$ the number of different orders on S.

LEMMA 4.2. *Let $(S, \prec)$ be a poset, and $S = U \sqcup V$ a partition of S. U and V are posets with the induced order.*

(a) Suppose that for all $u \in U$ and $v \in V$, $u \prec v$. Then

$$o(S) = o(U)o(V).$$

(b) Suppose that for all $u \in U$ and $v \in V$, u and v are unrelated to each other. Then

$$\frac{o(S)}{|S|!} = \frac{o(U)}{|U|!}\frac{o(V)}{|V|!}.$$

Proof. Part (a) is obvious. It is also clear that under the assumptions of part (b), there is a bijection between the orders on S and triples

$$\{\text{order on } U, \text{ order on } V, \text{ a subset of } \{1,2,\ldots,|S|\} \text{ of cardinality } |U|\}.$$

Therefore

$$o(S) = \binom{|S|}{|U|} o(U) o(V).$$

This implies part (b). $\square$

DEFINITION 4.3. For a non-crossing partition $\pi = \{V_1, V_2, \ldots, V_k\}$, define a partial order on it as follows: for $U, V \in \pi$, $U \prec V$ if for some $i, j \in U$ and any $v \in V$, we have $i < v < j$. In this case we say that U *covers* V. Minimal elements with respect to this order are called the *outer* blocks of π; the rest are the *inner* blocks.

See [HS11, HS14] for more on orders on non-crossing partitions.

DEFINITION 4.4. Let $\mu : \mathcal{B}\langle X\rangle \to \mathcal{B}$ be a $\mathcal{B}$-bimodule map; at this point no positivity assumptions are made. Its *monotone cumulant functional* is the $\mathcal{B}$-bimodule map $K^\mu : \mathcal{B}_0\langle X\rangle \to \mathcal{B}$ defined implicitly by

$$\mu[b_0 X b_1 X \ldots b_{n-1} X b_n] = \sum_{\pi \in \mathrm{NC}(n)} \frac{o(\pi)}{|\pi|!} K^\mu_\pi [b_0 X b_1 X \ldots b_{n-1} X b_n]. \tag{42}$$

Here for a non-crossing partition π, K^μ_π is defined in terms of K^μ in the usual way as in [Spe98] (see Section 3 of [ABFN13] for a detailed discussion), and $o(\pi)$ is the number of orders on π considered as a poset (as in the preceding definition). The implicit definition determines the monotone cumulants uniquely since

$$K^\mu[b_0 X \ldots b_{n-1} X b_n] = \mu[b_0 X \ldots b_{n-1} X b_n] - \sum_{\substack{\pi \in \mathrm{NC}(n) \\ \pi \neq \hat{1}_n}} \frac{o(\pi)}{|\pi|!} K^\mu_\pi [b_0 X \ldots b_{n-1} X b_n], \tag{43}$$

and the second term on the right-hand side can be expressed in terms of lower-order moments.

REMARK 4.5. For $N \in \mathbb{N}$, we note that

$$K^{\mu \otimes 1_N} = K^\mu \otimes 1_N.$$

The proof of this fact is identical to that of Proposition 6.3 of [PV13].
It follows that the generating function arguments in the rest of this section work equally well for each $\mu \otimes 1_N$, and so the corresponding generating functions completely determine the states.

LEMMA 4.6. *For $\mathcal{B}$-bimodule maps, $\mu_i \to \mu$ if and only if $K^{\mu_i} \to K^\mu$.*

Proof. By assumption, $\mu_i[b] = b = \mu[b]$. For $n \geq 1$, clearly if

$$K^{\mu_i}[b_0 X b_1 X \ldots b_{n-1} X b_n] \to K^\mu[b_0 X b_1 X \ldots b_{n-1} X b_n]$$

then

$$\mu_i[b_0 X b_1 X \ldots b_{n-1} X b_n] \to \mu[b_0 X b_1 X \ldots b_{n-1} X b_n]$$

from equation (42). The other implication follows by induction on n, using equation (43). $\square$

DEFINITION 4.7. For μ as above and $\eta : \mathcal{B} \to \mathcal{B}$ a linear map, define $\mu^{\rhd \eta}$ via

$$K^{\mu^{\rhd \eta}}[b_0 X b_1 X \ldots b_{n-1} X b_n] = b_0 \eta \left(K^{\mu}[X b_1 X \ldots b_{n-1} X] \right) b_n.$$

Define the formal generating functions

$$H^{\mu}(b) = \sum_{n=0}^{\infty} \mu[b(Xb)^n]$$

and

$$K^{\mu}(b) = \sum_{n=1}^{\infty} K^{\mu}[b(Xb)^n].$$

Note that as formal series,

$$H^{\mu}(b) = G^{\mu}(b^{-1}),$$

so our notation is consistent with the analytic function notation in the rest of the article, except that we use superscripts for formal series. Note also that these generating functions differ by a factor of b from the more standard ones, and are more appropriate for the computations with monotone convolution.

REMARK 4.8. Fix $n \in \mathbb{N}$ and $\pi \in \mathrm{NC}(n)$. Denote by $V_1, \ldots, V_k$ the outer blocks of π, by $c(V_i)$ the partition consisting of V_i and the inner blocks it covers, and by $c_j(V_i)$, $j = 1, 2, \ldots, |V_i| - 1$ the partition consisting of the inner blocks lying between the jth and the $(j+1)$st elements of V_i. By Lemma 4.2 part (b),

$$\frac{o(\pi)}{|\pi|!} = \prod_{i=1}^{k} \frac{o(c(V_i))}{|c(V_i)|!}. \tag{44}$$

By part (a) of that lemma,

$$o(c(V_i)) = o(\{V_i\}) o \left(\bigcup_{j=1}^{|V_i|-1} c_j(V_i) \right) = o \left(\bigcup_{j=1}^{|V_i|-1} c_j(V_i) \right)$$

and so by part (b),

$$\frac{o(c(V_i))}{(|c(V_i)|-1)!} = \prod_{j=1}^{|V_i|-1} \frac{o(c_j(V_i))}{|c_j(V_i)|!}. \tag{45}$$

The following results may be contained in [Pop08], and are closely related to Proposition 3.5 in [HS14]. We provide a purely combinatorial direct proof.

PROPOSITION 4.9. *Let $\mu : \mathcal{B}\langle X \rangle \to \mathcal{B}$ be an exponentially bounded $\mathcal{B}$-bimodule map. Then for each n*

$$\frac{dH^{(\mu \otimes 1_N)^{\rhd t}}(b)}{dt} = K^{\mu \otimes 1_N}(H^{(\mu \otimes 1_N)^{\rhd t}}(b)).$$

Proof. It suffices to prove the result for $N = 1$. We begin by proving this equality for each of the coefficients of the series expansions of $H^{\mu^{\rhd t}}$ and $K^\mu \circ H^{\mu^{\rhd t}}$. Since

$$\begin{aligned} \frac{d}{dt}\mu^{\rhd t}[b(Xb)^n] &= \frac{d}{dt}\sum_{\pi\in \mathrm{NC}(n)} t^{|\pi|}\frac{o(\pi)}{|\pi|!}K^\mu_\pi[bXbX\ldots bXb] \\ &= \sum_{\pi\in \mathrm{NC}(n)} t^{|\pi|-1}\frac{o(\pi)}{(|\pi|-1)!}K^\mu_\pi[bXbX\ldots bXb], \end{aligned} \tag{46}$$

the coefficient of $K^\mu_\pi[b(Xb)^n]$ in its expansion is $t^{|\pi|-1}\frac{o(\pi)}{(|\pi|-1)!}$. On the other hand,

$$\begin{aligned} &K^\mu\left[H^{\mu^{\rhd t}}(b)\left(XH^{\mu^{\rhd t}}(b)\right)^l\right] \\ &= K^\mu\left[H^{\mu^{\rhd t}}(b)XH^{\mu^{\rhd t}}(b)X\ldots H^{\mu^{\rhd t}}(b)XH^{\mu^{\rhd t}}(b)\right] \\ &= \sum_{k_0,\ldots,k_l\geq 0} K^\mu\Bigg[\sum_{\pi_0\in \mathrm{NC}(k_0)} t^{|\pi_0|}\frac{o(\pi_0)}{|\pi_0|!}K^\mu_{\pi_0}X \\ &\qquad \sum_{\pi_1\in \mathrm{NC}(k_1)} t^{|\pi_1|}\frac{o(\pi_1)}{|\pi_1|!}K^\mu_{\pi_1}X\ldots X\sum_{\pi_l\in \mathrm{NC}(k_l)} t^{|\pi_l|}\frac{o(\pi_l)}{|\pi_l|!}K^\mu_{\pi_l}\Bigg] \\ &= \sum_{k_0,\ldots,k_l\geq 0}\sum_{\substack{\pi_i\in \mathrm{NC}(k_i),\\ 0\leq i\leq l}}\frac{o(\pi_0)}{|\pi_0|!}\frac{o(\pi_1)}{|\pi_1|!}\cdots\frac{o(\pi_l)}{|\pi_l|!}K^\mu \\ &\qquad \left[K^\mu_{\pi_0}XK^\mu_{\pi_1}X\ldots XK^\mu_{\pi_l}\right]t^{|\pi_0|+|\pi_1|+\ldots+|\pi_l|}, \end{aligned}$$

where $K_\emptyset(b) = b$. Fixing $n = k_0 + \ldots + k_l + l$, each term in this expansion is a multiple of $K^\mu_\pi[b(Xb)^n]$, where π is constructed from partitions $\pi_0, \pi_1, \ldots, \pi_k$ and an additional outer block of l elements:

$$V = \{k_0 + 1, k_0 + k_1 + 2, \ldots, k_0 + \ldots + k_{l-1} + l\} \in \pi$$

and

$$\pi_i = \text{ restriction of } \pi \text{ to } [k_0 + \ldots + k_{i-1} + i + 1, k_0 + \ldots + k_i + i], \quad i = 0, 1, \ldots, l.$$

Note that $|\pi_0| + |\pi_1| + \ldots + |\pi_l| = |\pi| - 1$. This identification has an inverse, which requires first choosing one of the k outer blocks of π. Order the outer blocks left-to-right and call the specially chosen block V_i. Using the notation from Remark 4.8, we see that the coefficient of $K^\mu_\pi[b(Xb)^n]$ in the expansion of $K^\mu(H^{\mu^{\rhd t}}(b))$ is $t^{|\pi|-1}$

times

$$\sum_{i=1}^{k} \frac{o\left(\bigcup_{j<i} c(V_j)\right)}{\left|\bigcup_{j<i} c(V_j)\right|!} \left(\prod_{j=1}^{|V_i|-1} \frac{o(c_j(V_i))}{|c_j(V_i)|!}\right) \frac{o\left(\bigcup_{j>i} c(V_j)\right)}{\left|\bigcup_{j>i} c(V_j)\right|!}$$
$$= \sum_{i=1}^{k} \frac{o(c(V_i))}{(|c(V_i)|-1)!} \prod_{j\neq i} \frac{o(c(V_j))}{|c(V_j)|!}$$
$$= \prod_{j=1}^{k} \frac{o(c(V_j))}{|c(V_j)|!} \sum_{i=1}^{k} |c(V_i)|$$
$$= |\pi| \prod_{j=1}^{k} \frac{o(c(V_j))}{|c(V_j)|!}$$
$$= \frac{o(\pi)}{(|\pi|-1)!}.$$

Here we used equation (45), and equation (44) applied to partitions $\bigcup_{j<i} c(V_j)$ and $\bigcup_{j>i} c(V_j)$, in the first line, and again (44) in the last line. Since we obtained the same coefficient as in expansion (46), the result is proved for each of the individual components of the respective series expansions for each $n \in \mathbb{N}$.

Extending this to the series expansions and, therefore, the functions, observe that all of the sets over which the sums occur have cardinality whose growth rate is exponential over n. Thus, for $\|b\|$ small enough, the exponential boundedness of μ implies that the respective series are absolutely convergent. We may therefore conclude that the t coefficients of the series expansions agree, provided that $b \in B_\delta(0)$ for $\delta > 0$ small enough. Thus,

$$\frac{dH^{\mu^{\rhd t}}(b)}{dt} = K^\mu(H^{\mu^{\rhd t}}(b)).$$

for $b \in B_\delta(0)$.

To extend to arbitrary bounded sets in $\mathcal{B}^-$, consider the net of difference quotients

$$D_h^\mu(b,t) = \frac{H^{\mu^{\rhd t+h}}(b) - H^{\mu^{\rhd t}}(b)}{h}$$

for $t > 0$. We have just shown that

$$\lim_{h\to 0} D_h^\mu(b,t) \to K^\mu(H^{\mu^{\rhd t}}(b))$$

uniformly on $B_\delta(0)$. By Theorem 2.10 in [BPV12], this implies that the same is true on all bounded sets in $\mathcal{B}^-$. Thus, at the level of functions,

$$\frac{dH^{\mu^{\rhd t}}(b)}{dt} = K^\mu(H^{\mu^{\rhd t}}(b)),$$

proving our result. □

COROLLARY 4.10.
$$H^{(\mu\otimes 1_n)^{\rhd(s+t)}}(b) = H^{(\mu\otimes 1_n)^{\rhd s}}\left(H^{(\mu\otimes 1_n)^{\rhd t}}(b)\right).$$

In particular,

$$F^{\mu^{\rhd(s+t)}}(b) = F^{\mu^{\rhd s}}\left(F^{\mu^{\rhd t}}(b)\right),$$

so the combinatorial definition of monotone convolution powers coincides with the complex analytic one in Definition 2.3.

Proof. By Proposition 4.9, $H^{\mu^{\rhd s}}\left(H^{\mu^{\rhd t}}(b)\right)$, as a function of s, satisfies

$$\frac{d}{ds}H^{\mu^{\rhd s}}\left(H^{\mu^{\rhd t}}(b)\right) = K^{\mu}\left(H^{\mu^{\rhd s}}\left(H^{\mu^{\rhd t}}(b)\right)\right),$$

$$H^{\mu^{\rhd s}}\left(H^{\mu^{\rhd t}}(b)\right)\Big|_{s=0} = H^{\mu^{\rhd t}}(b).$$

Since, by the same proposition, $H^{\mu^{\rhd(s+t)}}(b)$ also satisfies this differential equation with this initial condition, they coincide for all positive s.
For the second statement, we observe that

$$G^{\mu^{\rhd s}}\left(F^{\mu^{\rhd t}}(b)\right) = G^{\mu^{\rhd s}}\left(\left(G^{\mu^{\rhd t}}(b)\right)^{-1}\right) = H^{\mu^{\rhd s}}\left(H^{\mu^{\rhd t}}(b^{-1})\right) =$$

$$= H^{\mu^{\rhd(s+t)}}(b^{-1}) = G^{\mu^{\rhd(s+t)}}(b).$$

□

Proposition 4.11. *If $\mu, \nu \in \Sigma_0$ and $\mu \rhd \mu = \nu \rhd \nu$, then $\mu = \nu$. In particular, if the square root with respect to the monotone convolution exists, it is unique.*

Proof. Under the given assumption,

$$K^{\mu} = \frac{1}{2}K^{\mu\rhd\mu} = K^{\nu},$$

and therefore $\mu = \nu$. □

Remark 4.12. Let $\gamma \in \mathcal{B}$ be self-adjoint, and $\sigma : \mathcal{B}\langle X\rangle \to \mathcal{B}$ be a completely positive but *not* necessarily a $\mathcal{B}$-bimodule map. Define $\nu_{\uplus}^{\gamma,\sigma}$ via its Boolean cumulant functional

$$B^{\nu_{\uplus}^{\gamma,\sigma}}[b_0Xb_1] = b_0\gamma b_1, \quad B^{\nu_{\uplus}^{\gamma,\sigma}}[b_0Xb_1X\ldots b_{n-1}Xb_n] = b_0\sigma[b_1X\ldots b_{n-1}]b_n.$$

It is known [BPV12, ABFN13] that $\nu_{\uplus}^{\gamma,\sigma}$ is a completely positive $\mathcal{B}$-bimodule map. Similarly, define $\nu_{\rhd}^{\gamma,\sigma}$ via its monotone cumulant functional

$$K^{\nu_{\rhd}^{\gamma,\sigma}}[b_0Xb_1] = b_0\gamma b_1, \quad K^{\nu_{\rhd}^{\gamma,\sigma}}[b_0Xb_1X\ldots b_{n-1}Xb_n] = b_0\sigma[b_1X\ldots b_{n-1}]b_n.$$

We could also define $\nu_{\boxplus}^{\gamma,\sigma}$ via its free cumulant functional

$$R^{\nu_{\boxplus}^{\gamma,\sigma}}[b_0Xb_1] = b_0\gamma b_1, \quad R^{\nu_{\boxplus}^{\gamma,\sigma}}[b_0Xb_1X\ldots b_{n-1}Xb_n] = b_0\sigma[b_1X\ldots b_{n-1}]b_n.$$

Lemma 4.13. *Let $k_i \to \infty$ be a numerical sequence, $\{\mu_i : \mathcal{B}\langle X\rangle \to \mathcal{B}\}_{i=1}^{\infty}$ a sequence of linear $\mathcal{B}$-bimodule maps, and $\rho : \mathcal{B}_0\langle X\rangle \to \mathcal{B}$ a linear $\mathcal{B}$-bimodule map. The following are equivalent.*

(a) $k_i\mu_i[P(X)] \to \rho[P(X)]$ *for all* $P(X) \in \mathcal{B}_0\langle X\rangle$.
(b) $k_iR^{\mu_i}[P(X)] \to \rho[P(X)]$ *for all* $P(X) \in \mathcal{B}_0\langle X\rangle$.
(c) $k_iB^{\mu_i}[P(X)] \to \rho[P(X)]$ *for all* $P(X) \in \mathcal{B}_0\langle X\rangle$.

(d) $k_i K^{\mu_i}[P(X)] \to \rho[P(X)]$ *for all* $P(X) \in \mathcal{B}_0\langle X\rangle$.

Here in all cases, the convergence is in norm on $\mathcal{B}$.

Proof. We will prove the equivalence between (a) and (d); the rest are similar, and were proved in [BPV12]. Indeed, on $\mathcal{B}_0\langle X\rangle$,

$$k_i\mu_i[b_0Xb_1X\ldots b_{n-1}Xb_n] = k_iK^{\mu_i}[b_0Xb_1X\ldots b_{n-1}Xb_n] + \sum_{\substack{\pi\in \mathrm{NC}(n)\\ |\pi|\geq 2}} \frac{1}{k_i^{|\pi|-1}}\frac{o(\pi)}{|\pi|!}\,(k_iK^{\mu_i})_\pi\,[b_0Xb_1X\ldots b_{n-1}Xb_n].$$

It follows immediately that (d) implies (a). The converse implication follows by induction on n. □

COROLLARY 4.14. *For linear* $\mathcal{B}$*-bimodule maps* $\mu_i : \mathcal{B}\langle X\rangle \to \mathcal{B}$*, the following are equivalent.*

(a)
$$k_i\mu_i[X] \to \gamma, \quad k_i\mu_i[Xb_1X\ldots b_{n-1}X] \to \sigma[b_1X\ldots b_{n-1}].$$

(b)
$$\mu_i^{\boxplus k_i} \to \nu_{\boxplus}^{\gamma,\sigma}.$$

(c)
$$\mu_i^{\uplus k_i} \to \nu_{\uplus}^{\gamma,\sigma}.$$

(d)
$$\mu_i^{\rhd k_i} \to \nu_{\rhd}^{\gamma,\sigma}.$$

Proof. We will prove the equivalence between (a) and (d); the rest are similar, see Lecture 13 in [NS06]. Indeed, by Lemma 4.6, the statement in part (d) is equivalent to

$$k_iK^{\mu_i} \to K^{\nu_{\rhd}^{\gamma,\sigma}},$$

which by definition of $\nu_{\rhd}^{\gamma,\sigma}$ means

$$k_iK^{\mu_i}[X] \to \gamma, \quad k_iK^{\mu_i}[Xb_1X\ldots b_{n-1}X] \to \sigma[b_1X\ldots b_{n-1}]$$

This is equivalent to (a) by the preceding lemma. □

COROLLARY 4.15. $\nu_{\rhd}^{\gamma,\sigma}$ *is a completely positive map.*

Proof. We can choose completely positive μ_i such that $\mu_i^{\uplus i} \to \nu_{\uplus}^{\gamma,\sigma}$, for example by taking $\mu_i = \nu_{\uplus}^{\frac{1}{i}\gamma,\frac{1}{i}\sigma}$. Then $\nu_{\rhd}^{\gamma,\sigma}$ is the limit of completely positive maps $\mu_i^{\rhd i}$, and as such is completely positive (monotone convolution of two completely positive maps is known to be positive, see Proposition 6.2 of [Pop08] and also [Pop12]). □

PROPOSITION 4.16. *Monotone convolution semigroups of completely positive* $\mathcal{B}$*-bimodule maps are in a one-to-one correspondence with pairs* (γ,σ) *as above.*

Proof. $\left\{\nu_{\triangleright}^{t\gamma,t\sigma} : t \geq 0\right\}$ form a one-parameter monotone convolution semigroup of completely positive $\mathcal{B}$-bimodule maps. Conversely, if $\{\mu_t\}$ is such a semigroup, define

$$\gamma = \left.\frac{d}{dt}\right|_{t=0} \mu_t[X] = K^{\mu_1}[X] \in \mathcal{B}^{sa},$$

$$\sigma[b_1 X \dots b_{n-1}] = \left.\frac{d}{dt}\right|_{t=0} \mu_t[X b_1 X \dots b_{n-1} X] = K^{\mu_1}[X b_1 X \dots b_{n-1} X].$$

Since for $P_i \in \mathcal{B}\langle X\rangle$ and $c_i \in \mathcal{B}$,

$$\sum_{i,j=1}^{N} c_i^* \sigma[P_i^* P_j] c_j = \left.\frac{d}{dt}\right|_{t=0} \mu_t \left[\sum_{i,j=1}^{N} c_i^* X P_i^* P_j X c_j\right] =$$

$$= \lim_{t \downarrow 0} \frac{1}{t} \mu_t \left[\sum_{i,j=1}^{N} c_i^* X P_i^* P_j X c_j\right] \geq 0,$$

σ is completely positive $\square$

Remark 4.17. A short calculation shows that

$$\Phi(b) = \gamma + G_\sigma(b).$$

This, combined with Theorem 2.1, gives an alternative proof of the result in Theorem 3.5 that generators of semigroups arising from $\triangleright$-infinitely divisible distributions coincide with the set Λ. One can also use a standard combinatorial argument to show that $\triangleright$-infinitely divisible distributions belong to such one-parameter semigroups. At this point, we do not know how to obtain the more general results in Theorem 3.5 by combinatorial methods.

Appendix A. Characterization of general Cauchy transforms

In this appendix, we extend the main result in [Wil13], namely the classification of the Cauchy transforms associated to distributions $\mu \in \Sigma_0$, to the Cauchy transforms associated to more general CP maps.

Theorem A.1. The following are equivalent:

(I) The analytic non-commutative function $G = (G^{(n)})_{n\geq 1} : H^+(\mathcal{B}) \to H^-(\mathcal{B})$ has the property that $H = (H^{(n)})_{n\geq 1}$ defined through the equalities $H^{(n)}(b) := G^{(n)}(b^{-1})$ for all $n \in \mathbb{N}$ and $b \in M_n(\mathcal{B})$ has uniformly analytic extension to a neighborhood of 0 satisfying $H^{(n)}(0) = 0$.

(II) There exists a $\mathbb{C}$-linear map $\sigma : \mathcal{B}\langle X\rangle \to \mathcal{B}$ satisfying (1) and (2) such that $G^{(n)}(b) = \sigma((b - X)^{-1})$.

Proof. We begin with (II) $\Rightarrow$ (I). Let σ satisfy (1) and (2). By [PV13], Lemma 5.8, we may conclude that there exists a $\boxplus$-infinitely divisible distribution $\mu \in \Sigma_0$ such that $\rho_\mu(XP(X)X) = \sigma(P(X))$ for all $P(X) \in \mathcal{B}\langle X\rangle$ (here, ρ_μ denotes the free cumulant function associated to μ). Thus, the Voiculescu transform of μ satisfies the following equality:

$$\varphi_\mu^{(n)}(b) = -\sigma((b - X)^{-1}) \tag{47}$$

for all $n \in \mathbb{N}$ and where the inverse in the equality is considered as a geometric series, so that the right hand side is convergent for $\|b^{-1}\|$ small enough dependent on (1). Since μ is $\boxplus$-infinitely divisible, by Proposition 5.1 in [Wil13], we have that the left hand side of (47) extends to

$$H^+(\mathcal{B}) \cup H^-(\mathcal{B}) \bigcup_{n=1}^{\infty} \{b \in M_n(\mathcal{B}) : \|b^{-1}\| < C\}$$

where C is a fixed constant, independent of n.
Now, by Proposition 1.2 in [PV13], the fact that $\mu \in \Sigma_0$ implies that μ is realized as the distribution arising from a non-commutative probability space $(\mathcal{A}, E, \mathcal{B})$. That is,

$$\mu(P(X)) = E(P(a))$$

for a fixed self-adjoint element $a \in \mathcal{B}$ and all $P(X) \in \mathcal{B}\langle X \rangle$. Thus, $\sigma((b - X)^{-1}) = \rho_\mu(a(b - a)^{-1}a)$ and, since $b - a \in M_n^+(\mathcal{B})$ and ρ_μ is a CP map on $\mathcal{B}\langle X \rangle_0$ we may conclude that the $\sigma((b - X)^{-1}) \in M_n^-(\mathcal{B})$ for all $b \in M_n^+(\mathcal{B})$.
Further note that

$$H(b) = \sigma((b^{-1} - X)^{-1}) = \sum_{k=0}^{\infty} \sigma((bX)^k b)$$

is convergent in a neighborhood of zero since σ satisfies (1). It is also immediate that $H(0) = 0$. This completes one direction of our proof.
We now prove (I) $\Rightarrow$ (II). We will follow the proof of Theorem 4.1 in [Wil13] and refer to this paper for the appropriate terminology.
We recover our operator σ through the differential structure of H. Indeed, we define the map σ by letting

$$(\sigma \otimes 1_n)(b_1(X \otimes 1_n)b_2 \cdots (X \otimes 1_n)b_{\ell+1}) := \Delta_{\mathcal{R}}^{\ell+1} H^{(n)}(\underbrace{0, \ldots, 0}_{\ell+2 \,-\, \text{times}})(b_1, b_2, \ldots, b_{\ell+1})$$

for elements $b_1, b_2, \cdots, b_{\ell+1} \in M_n(\mathcal{B})$. It is a consequence of Proposition 3.1 in [Wil13] and [KVV14], Theorem 3.10 that this is a well defined operator. Moreover, the equality

$$\Delta_{\mathcal{R}}^{\ell+1} H^{(n)}(\underbrace{0, \ldots, 0}_{\ell+2 \,-\, \text{times}})(b, b, \ldots, b) = \frac{1}{(\ell+1)!} \frac{d^{\ell+1}}{dt^{\ell+1}} H^{(n)}(0 + tb)|_{t=0}$$

and the fact that the function is analytic in a neighborhood of 0 implies that

$$H^{(n)}(b) = \sum_{k=0}^{\infty} (\sigma \otimes 1_n)((bX)^k b) \tag{48}$$

once we show that σ satisfies (1). Continuation will allow us to conclude that

$$G^{(n)}(b) = \sum_{k=0}^{\infty} (\sigma \otimes 1_n)((b^{-1}X)^k b^{-1}) = (\sigma \otimes 1_n)((b - X)^{-1}). \tag{49}$$

Thus, our theorem will follow when we can show that σ satisfies properties (1) and (2).

To prove (1), we note that this is equivalent to showing that

$$\|\sigma(b_1 X b_2 \cdots X b_{\ell+1})\| \leq C M^{\ell+1}$$

for a fixed $C > 0$, provided that $\|b_1\| = \cdots = \|b_{\ell+1}\| = 1$. This will follow from uniform analyticity and matches the proof of the same fact in [Wil13]. Indeed, consider the element of $M_{\ell+2}(\mathcal{B})$

$$B = \begin{pmatrix} 0 & b_1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & b_2 & 0 & \cdots & 0 \\ 0 & 0 & 0 & b_3 & \cdots & 0 \\ & \vdots & & & \vdots & \\ 0 & 0 & 0 & 0 & \cdots & b_{\ell+1} \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Note that $H^{(\ell+1)}$ has a bound of C on a ball of radius r about 0, independent of ℓ since we are assuming that H is uniformly analytic. Thus,

$$\begin{aligned}\|\sigma(b_1 X b_2 \cdots X b_{\ell+1})\| &= \frac{\|\delta^{\ell+1} H^{(\ell+2)}(0; B)\|}{(\ell+1)!} \\ &= \|\Delta_{\mathcal{R}}^{\ell+1} H^{(\ell+2)}(0,\ldots,0)(B,\ldots,B)\| \\ &= \|r^{-(\ell+1)} \Delta_{\mathcal{R}}^{\ell+1} H^{(\ell+2)}(0,\ldots,0)(rB,\ldots,rB)\| \\ &= \left(\frac{1}{r}\right)^{\ell+1} \frac{\|\delta^{\ell+1} H^{(\ell+2)}(0; rB)\|}{(\ell+1)!} \\ &\leq C \left(\frac{1}{r}\right)^{\ell+1}\end{aligned}$$

where the last inequality follows from the Cauchy estimates in Theorem 2.2.

We must prove the technical fact that fact that

$$\sigma|_{M_n(\mathcal{B})} \geq 0 \tag{50}$$

Assume that $\sigma(P) < 0$ for some $P \in M_n^+(\mathcal{B})$ where we can assume that $P > \delta 1$ for some $\delta > 0$. Note that $G^{(n)}(zP^{-1}) \in M_n^-(\mathcal{B})$ for all $z \in \mathbb{C}^+$ by assumption so that $\lambda G^{(n)}(i\lambda P^{-1}) \in M_n^-(\mathcal{B})$ for all $\lambda \in \mathbb{R}^+$. Utilizing the series expansion in (49) as well as the exponential bound that we have just proven, we conclude that the

$$\lim_{\lambda \uparrow \infty} \lambda G^{(n)}(i\lambda P^{-1}) = \frac{\sigma(P)}{i} = -i\sigma(P) \notin M_n^-(\mathcal{B}).$$

This contradiction implies (50).

It remains to show (2). Once again, this will closely follow the proof of the analogous fact in Theorem 4.1 in [Wil13]. Indeed, we will first show that

$$(\sigma \otimes 1_n)(P(X \otimes 1_n + b_0)^* P(X \otimes 1_n + b_0)) \geq 0 \tag{51}$$

for any monomial $P(X) = b_1(X \otimes 1_n) b_2 \cdots X \otimes 1_n b_{\ell+1} \in M_n(\mathcal{B})\langle X \rangle$ and $b_0 \in M_n(\mathcal{B})$. We also assume that $|b_{\ell+1}| > \epsilon 1_n$ and the general case follows by letting $\epsilon \downarrow 0$.

Towards this end, we consider elements $C, E_0, E_1 \in M_{n(\ell+1)}(\mathcal{B})$ defined as follows:

$$C = \begin{pmatrix} 0 & c_1 & 0 & 0 & 0 & \cdots & 0 \\ c_1^* & 0 & c_2 & 0 & 0 & \cdots & 0 \\ 0 & c_2^* & 0 & c_3 & 0 & \cdots & 0 \\ \vdots & & & \vdots & & & \vdots \\ 0 & 0 & \cdots & 0 & c_{\ell-1}^* & 0 & c_\ell \\ 0 & 0 & \cdots & 0 & 0 & c_\ell^* & |c_{\ell+1}|^2 \end{pmatrix}; E_0 = \underbrace{1_n \oplus 1_n \oplus \cdots \oplus 1_n}_{\ell \ times} \oplus 0_n$$

and $E_1 = 1_{n(\ell+1)} - E_0$ where $c_i = \delta b_i$ for $i = 1, \ldots, \ell$ and $c_{\ell+1} = b_{\ell+1}/\delta^\ell$ for $\delta > 0$ to be specified. Note that $b_1 X b_2 \cdots X b_{\ell+1} = c_1 X c_2 \cdots X c_{\ell+1}$. We define a function

$$\hat{g}^{n(\ell+1)}(b) := G^{n(\ell+1)}(b - b_0) : M^+_{n(\ell+1)}(\mathcal{B}) \to M^-_{n(\ell+1)}(\mathcal{B})$$

The following properties are rather trivial and their proof matches those of Theorem 4.1 in [Wil13].

(a) $C + \epsilon E_0 > \gamma 1_n$ for some $\gamma > 0$ provided that $\delta > 0$ is small enough.
(b) The $n \times n$ minor in the top left corner of

$$[(C + \epsilon E_0)(X \otimes 1_{n(\ell+1)} + b_0 \otimes 1_{\ell+1})]^{2(\ell-1)}(C + \epsilon E_0)$$

is equal to $P(X + b_0)P^*(X + b_0) + O(\epsilon)$.
(c) $\hat{g}^{(n(\ell+1))}(b) = \sum_{p=0}^\infty \sigma([b^{-1}(X \otimes 1_{n(\ell+1)} + b_0 \otimes 1_{\ell+1})]^p b^{-1})$ for b^{-1} in a neighborhood of 0.
(d) We have that $z\hat{g}^{(n(\ell+1))}(zb) \to \sigma(b^{-1})$ in norm as $|z| \uparrow \infty$ for $b > \gamma 1_n$.
(e) $\hat{h}^{(n(\ell+1))}(b) := \hat{g}^{(n(\ell+1))}(b^{-1})$ has analytic extension to a neighborhood of zero.

The only one of these properties that differs from the proof of Theorem 4.1 in [Wil13] is (d). It follows immediately from the series expansion in (48).

We now have the pieces in place to prove (51). Note that (a) implies that $C + \epsilon E_0$ is invertible so that the map

$$z \mapsto \hat{g}^{(n(\ell+1))}(z(C + \epsilon E_0)^{-1})$$

sends $\mathbb{C}^+$ into $M_n(\mathcal{B})^-$. Let $B_{i,j} \in M_n(\mathcal{B})$ for $i, j = 1, \ldots, \ell+1$ and consider the element $B = (B_{i,j})_{i,j=1}^{\ell+1} \in M_{n(\ell+1)}(\mathcal{B})$. Given a state $f \in M_n(\mathcal{B})^*$ we define a new state

$$f_{1,1}(B) := f(B_{1,1}) : M_{n(\ell+1)}(B) \to \mathbb{C}.$$

We may define a map

$$G_{f,C,\epsilon}(z) = f_{1,1} \circ \hat{g}^{(n(\ell+1))}(z(C + \epsilon E_0)^{-1}) : \mathbb{C}^+ \to \mathbb{C}^-.$$

Properties (c) and (d) imply the following for $z \in \mathbb{C}^+$:

$$\begin{aligned} \lim_{|z|\uparrow\infty} zG_{f,C,\epsilon}(z) &= \lim_{|z|\uparrow\infty} f_{1,1}\left[z\hat{g}^{(n(\ell+1))}(z(C + \epsilon E_0)^{-1})\right] \\ &= f_{1,1}(\sigma(C + \epsilon E_0)) \geq 0 \end{aligned}$$

where the last inequality will follow from the fact that $f_{1,1}$ is a state, property (a) and (50).

Now, observe that the coefficient of $z^{-2\ell+1}$ in the function $G_{f,C,\epsilon}$ is equal to $\rho(t^{2(\ell-1)}) > 0$. Furthermore, since

$$G_{f,C,\epsilon}(z) = G_\rho(z) = \sum_{\ell=0}^{\infty} \frac{\rho(t^\ell)}{z^{\ell+1}}$$
$$= \sum_{\ell=0}^{\infty} \frac{f_{1,1}(\sigma([(C+\epsilon E_0)(X \otimes 1_{n(\ell+1)} + b_0)]^\ell (C+\epsilon E_0)))}{z^{\ell+1}}$$

we may conclude that

$$f_{1,1} \circ \sigma([(C+\epsilon E_0)(X \otimes 1_{n(\ell+1)} + b_0)]^{2(\ell-1)}(C+\epsilon E_0)) = \rho(t^{2(\ell-1)}) \geq 0.$$

Recalling (b), it follows that $f \circ \sigma([P(X+b_0)P^*(X+b_0) + O(\epsilon)]) \geq 0$. Letting $\epsilon \downarrow 0$ and noting that f was an arbitrary state, we have proven that

$$(\sigma \otimes 1_n)(P(X+b_0)P^*(X+b_0)) \geq 0$$

for any monomial $P(X) \in M_n(\mathcal{B})\langle X \rangle$.
The extension from the case of monomials to general elements in $\mathcal{B}\langle X \rangle$ follows the proof in [Wil13] exactly so we will refrain from repeating it. This implies (2) and, therefore, our theorem. $\square$

References

[ABFN13] Michael Anshelevich, Serban T. Belinschi, Maxime Fevrier, and Alexandru Nica, *Convolution powers in the operator-valued framework*, Trans. Amer. Math. Soc. 365 (2013), no. 4, 2063–2097. MR 3009653

[AW14] Michael Anshelevich and John D. Williams, *Limit theorems for monotonic convolution and the Chernoff product formula*, Int. Math. Res. Not. IMRN (2014), no. 11, 2990–3021. MR 3214313

[BPV12] Serban T. Belinschi, Mihai Popa, and Victor Vinnikov, *Infinite divisibility and a non-commutative Boolean-to-free Bercovici-Pata bijection*, J. Funct. Anal. 262 (2012), no. 1, 94–123. MR 2852257

[BP99] Hari Bercovici and Vittorino Pata, *Stable laws and domains of attraction in free probability theory*, Ann. of Math. (2) 149 (1999), no. 3, 1023–1060, With an appendix by Philippe Biane. MR 2000i:46061

[BPo78] Earl Berkson and Horacio Porta, *Semigroups of analytic functions and composition operators*, Michigan Math. J. 25 (1978), 101–115. MR 0480965 (58 #1112)

[Dun38] Nelson Dunford, *Uniformity in linear spaces*, Trans. Amer. Math. Soc. 44 (1938), no. 2, 305-356. MR 1501971

[HP74] Einar Hille and Ralph S. Phillips, *Functional analysis and semi-groups*, American Mathematical Society, Providence, R. I., 1974, Third printing of the revised edition of 1957, American Mathematical Society Colloquium Publications, Vol. XXXI. MR 0423094 (54 #11077)

[HS11] Takahiro Hasebe and Hayato Saigo, *The monotone cumulants*, Ann. Inst. Henri Poincaré Probab. Stat. 47 (2011), no. 4, 1160–1170. MR 2884229

[HS14] ______, *On operator-valued monotone independence*, Nagoya Math. J. 215 (2014), 151–167. MR 3263527

[KVV14] Dmitry S. Kaliuzhnyi-Verbovetskyi and Victor Vinnikov, *Foundations of free noncommutative function theory*, Mathematical Surveys and Monographs, vol. 199, American Mathematical Society, Providence, RI, 2014. MR 3244229

[Mur00] Naofumi Muraki, *Monotonic convolution and monotonic levy-hincin formula*, preprint, 2000.

[Mur01] ______, *Monotonic independence, monotonic central limit theorem and monotonic law of small numbers*, Infin. Dimens. Anal. Quantum Probab. Relat. Top. 4 (2001), no. 1, 39–58. MR 1824472 (2002e:46076)

[NS06] Alexandru Nica and Roland Speicher, *Lectures on the combinatorics of free probability*, London Mathematical Society Lecture Note Series, vol. 335, Cambridge University Press, Cambridge, 2006. MR MR2266879 (2008k:46198)

[Pop08] Mihai Popa, *A combinatorial approach to monotonic independence over a C^*-algebra*, Pacific J. Math. 237 (2008), no. 2, 299–325. MR 2421124 (2009k:46119)

[Pop12] ______, *Realization of conditionally monotone independence and monotone products of completely positive maps*, J. Operator Theory 68 (2012), no. 1, 257–274. MR 2966045

[PV13] Mihai Popa and Victor Vinnikov, *Non-commutative functions and the non-commutative free Lévy-Hinčin formula*, Adv. Math. 236 (2013), 131–157. MR 3019719

[Spe98] Roland Speicher, *Combinatorial theory of the free product with amalgamation and operator-valued free probability theory*, Mem. Amer. Math. Soc. 132 (1998), no. 627, x+88. MR 1407898 (98i:46071)

[Voi87] Dan-Virgil Voiculescu, *Multiplication of certain noncommuting random variables*, J. Operator Theory, 18 (1987), 223–235. MR 915507 (89b:46076)

[Tay72] Joseph L. Taylor, *A general framework for a multi-operator functional calculus*, Adv. Math. 9 (1972), 183–252. MR 0328625 (48 #6967)

[Tay73] ______, *Functions of several noncommuting variables*, Bull. Amer. Math. Soc. 79 (1973), 1-34. MR 0315446 (47 #3995)

[Wil13] John D. Williams, *Analytic function theory for operator-valued free probability*, `arXiv:1309.0877 [math.OA]`, to be published in Crelle's journal, 2013.

Michael Anshelevich
Department of Mathematics
Texas A&M
University
College Station
TX
77843-3368
manshel@math.tamu.edu

John D. Williams
Universität des Saarlandes
Fachrichtung Mathematik
Postfach 151150
66041.
Saarbrücken
williams@math.uni-sb.de

On Geometric Aspects of Diffuse Groups

Steffen Kionke, Jean Raimbault
With an Appendix by Nathan Dunfield

Received: December 4, 2014
Revised: October 26, 2015

Communicated by Ursula Hamenstädt

Abstract. Bowditch introduced the notion of diffuse groups as a geometric variation of the unique product property. We elaborate on various examples and non-examples, keeping the geometric point of view from Bowditch's paper. In particular, we discuss fundamental groups of flat and hyperbolic manifolds. Appendix B settles an open question by providing an example of a group which is diffuse but not left-orderable.

2010 Mathematics Subject Classification: Primary 22E40; Secondary 06F15, 57M07, 20F65
Keywords and Phrases: orderable groups, discrete subgroups of Lie groups, hyperbolic manifolds, Bieberbach groups

Contents

1. Introduction

Following B. Bowditch [10], we say that a group Γ is *diffuse* if every finite non-empty subset $A \subset \Gamma$ has an *extremal point*, that is, an element $a \in A$ such that for any $g \in \Gamma \setminus \{1\}$ either ga or $g^{-1}a$ is not in A (see also 2.1 below). A non-empty finite set without extremal points will be called a *ravel*[1]; thus a group is diffuse if and only if it does not contain a ravel. Every non-trivial finite subgroup of Γ is a ravel, hence a diffuse group is torsion-free. In this work, we use geometric methods to discuss various examples of diffuse and non-diffuse groups.

The interest in diffuse groups stems from Bowditch's observation that they have the unique product property (see Section 2.2 below). Originally, unique products were introduced in the study of group rings of discrete, torsion-free groups. More precisely, it is easily seen that if a group Γ has unique products, then it satisfies Kaplansky's unit conjecture. In simple terms, this means that the units in the group ring $\mathbb{C}[\Gamma]$ are all *trivial*, i.e. of the form λg with $\lambda \in \mathbb{C}^\times$ and $g \in \Gamma$. A similar question can be asked replacing $\mathbb{C}$ by some integral domain. A weaker conjecture (Kaplansky's zero divisor conjecture) asserts that $\mathbb{C}[\Gamma]$ contains no zero divisor, and a still weaker one asserts that it contains no idempotents other than 1_Γ. There are other approaches to the zero divisor and idempotent conjecture (see for example [5], [47, Chapter 10]) which have succeeded in proving it for large classes of groups, whereas the unit conjecture has (to the best of our knowledge) only been tackled by establishing the possibly stronger unique product property. Consequently it is still unknown if the unit conjecture holds, for example, for all torsion-free groups in the class of crystallographic groups (see [23] for more on the subject), while the zero-divisor conjecture is known to hold (among other) for all torsion-free groups in the finite-by-solvable class, as proven by Kropholler, Linnell and Moody in [45].

There are further applications of the unique product property. For instance, if Γ has unique products, then it satisfies a conjecture of Y. O. Hamidoune on the size of isoperimetric atoms (cf. Conjecture 10 in [7]). Let us also mention that it is known that torsion-free groups without unique products exist, see for instance [57],[54],[61],[3],[19]. We note that for the examples in [57] (and their generalization in [61]) it is not known if the zero-divisor conjecture holds.

Using Lazard's theory of analytic pro-p groups, one can show that every arithmetic group Γ has a finite index subgroup Γ' such that the group ring $\mathbb{Z}[\Gamma']$ satisfies the zero divisor conjecture. This work originated from the idea to study Kaplansky's unit conjecture *virtually*. In this spirit we establish virtual diffuseness for various classes of groups and, moreover, we discuss examples of diffuse and non-diffuse groups in order to clarify the border between the two. Our results are based on geometric considerations.

1.1. Results.

[1] We think of this as an entangled ball of string.

1.1.1. *Crystallographic groups.* The torsion-free crystallographic groups, also called Bieberbach groups, are virtually diffuse since free abelian groups are diffuse. However, already in dimension three there is a Bieberbach group Δ_P which is not diffuse [10]. In fact, Promislow even showed that the group Δ_P does not satisfy the unique product property [54]. On the other hand, the nine other 3-dimensional Bieberbach groups are diffuse. So is there an easy way to decide whether a given Bieberbach group is diffuse or not? In Section 3 we discuss this question and show that in many cases it suffices to know the holonomy group.

Theorem A. Let Γ be a Bieberbach group with holonomy group G.

(i) If G is not solvable, then Γ is not diffuse.
(ii) If G has only cyclic Sylow subgroups, then Γ is diffuse.

Note that a finite group G with cyclic Sylow subgroups is meta-cyclic, thus solvable. We further show that in the remaining case, where G is solvable and has a non-cyclic Sylow subgroup, the group G is indeed the holonomy of both a diffuse and a non-diffuse Bieberbach group. Moreover, we give a complete list of the 16 non-diffuse Bieberbach groups in dimension four. Our approach is based on the equivalence of diffuseness and local indicability for amenable groups as obtained by Linnell and Witte Morris [46]. We include a new geometric proof of their result for the special case of virtually abelian groups.

1.1.2. *Discrete subgroups of rank-one Lie groups.* The class of hyperbolic groups is one of the main sources of examples of diffuse groups in [10]: it is an immediate consequence of Corollary 5.2 loc. cit. that any residually finite word-hyperbolic group contains with finite index a diffuse subgroup (the same statement for unique products was proven earlier by T. Delzant [24]). In particular, cocompact discrete subgroups of rank one Lie groups are virtually diffuse (for example, given an arithmetic lattice Γ in such a Lie group, any normal congruence subgroup of Γ of sufficiently high level is diffuse). On the other hand, not much is known in this respect about relatively hyperbolic groups, and it is natural to ask whether a group which is hyperbolic relative to diffuse subgroups must itself be virtually diffuse. In this paper we answer this question in the affirmative in the case of non-uniform lattices in rank one Lie groups.

Theorem B. If Γ is a lattice in one of the Lie groups $\mathrm{SO}(n,1), \mathrm{SU}(n,1)$ or $\mathrm{Sp}(n,1)$ then there is a finite-index subgroup $\Gamma' \leq \Gamma$ such that Γ' is diffuse.

In the case of an arithmetic lattice, the proof actually shows that normal congruence subgroups of sufficiently large level are diffuse. We left open the case of non-uniform lattices in the exceptional rank one group F_4^{-20}, but it is almost certain that our proof adapts also to this case. Theorem B is obtained as a corollary of a result on a more general class of geometrically finite groups of isometries. Another consequence is the following theorem.

Theorem C. Let Γ be any discrete, finitely generated subgroup of $\mathrm{SL}_2(\mathbb{C})$. There exists a finite-index subgroup $\Gamma' \leq \Gamma$ such that Γ' is diffuse.

The proofs of these theorems use the same approach as Bowditch's, that is a metric criterion (Lemma 2.1 below) for the action on the relevant hyperbolic space. The main new point we have to establish concerns the behaviour of unipotent isometries: the result we need (Proposition 4.2 below) is fairly easy to observe for real hyperbolic spaces; for complex ones it follows from a theorem of M. Phillips [53], and we show that the argument used there can be generalized in a straightforward way to quaternionic hyperbolic spaces. We also study axial isometries of real hyperbolic spaces in some detail, and give an optimal criterion (Proposition 4.5) which may be of use in determining whether a given hyperbolic manifold has a diffuse fundamental group.

1.1.3. *Three–manifold groups.* Following the solution of both Thurston's Geometrization conjecture (by G. Perelman [51, 52]) and the Virtually Haken conjecture (by I. Agol [2] building on work of D. Wise) it is known by previous work of J. Howie [40], and S. Boyer, D. Rolfsen and B. Wiest [12] that the fundamental group of any compact three–manifold contains a left-orderable finite-index subgroup. Since left-orderable groups are diffuse (see Section 2.2 below) this implies the following.

Theorem D. Let M be a compact three–manifold, then there is a finite-index subgroup of $\pi_1(M)$ which is diffuse.

Actually, one does not need Agol's work to prove this weaker result: the case of irreducible manifolds with non-trivial JSJ-decomposition is dealt with in [12, Theorem 1.1(2)], and non-hyperbolic geometric manifolds are easily seen to be virtually orderable. Finally, closed hyperbolic manifolds can be handled by Bowditch's result (see (iv) in Section 2.1 below).

We give a more direct proof of Theorem D in Section 5; the tools we use (mainly a 'virtual' gluing lemma) may be of independent interest. The relation between diffuseness (or unique products) and left-orderability is not very clear at present; in Appendix B Nathan Dunfield gives an example of a compact hyperbolic three-manifold whose fundamental group is not left-orderable, but nonetheless diffuse.

Acknowledgements. We are pleased to thank to George Bergman, Andres Navas and Markus Steenbock for valuable comments on a first version of this paper. The second author would especially like to thank Pierre Will for directing him to the article [53]. We thank the anonymous referee for comments improving the exposition.

Both authors are grateful to the Max-Planck-Institut für Mathematik in Bonn, where this work was initially developed, and which supported them financially during this phase.

2. Diffuse groups

We briefly review various notions and works related to diffuseness and present some questions and related examples of groups.

2.1. A quick survey of Bowditch's paper. We give here a short recapitulation of some of the content in Bowditch's paper [10]. The general notion of a diffuse action of a group is introduced there and defined as follows: let Γ be a group acting on a set X. Given a finite subset $A \subset X$, an element $a \in A$ is said to be an *extremal point* in A, if for all $g \in \Gamma$ which do not stabilize a then either ga or $g^{-1}a$ is not in A. The action of Γ on X is said to be diffuse if every finite subset A of X with $|A| \geq 2$ has at least *two* extremal points. An action in which each finite subset has at least *one* extremal point is called weakly diffuse by Bowditch; we will not use this notion in the sequel. It was observed by Linnell and Witte-Morris [46, Prop.6.2.] that a free action is diffuse if and only if it is weakly diffuse. Thus a group is diffuse (in the sense given in the introduction) if and only if its action on itself by left-translations is diffuse. More generally, Bowditch proves that if a group admits a diffuse action whose stabilizers are diffuse groups, then the group itself is diffuse. In particular, an extension of diffuse groups is diffuse as well.

The above can be used to deduce the diffuseness of many groups. For example, strongly polycyclic groups are diffuse since they are, by definition, obtained from the trivial group by taking successive extensions by $\mathbb{Z}$. Bowditch's paper provides many more examples of diffuse groups:

(i) The fundamental group of a compact surface of nonpositive Euler characteristic is diffuse;
(ii) More generally, any free isometric action of a group on an $\mathbb{R}$-tree is diffuse;
(iii) A free product of two diffuse groups is itself diffuse;
(iv) A closed hyperbolic manifold with injectivity radius larger than $\log(1+\sqrt{2})$ has a diffuse fundamental group.

We conclude this section with the following simple useful lemma, which appears as Lemma 5.1 in [10].

Lemma 2.1. *If Γ acts on a metric space (X, d_X) satisfying the condition*

$$(*) \qquad \forall x, y \in X, g \in \Gamma : \; gx \neq x \implies \max(d_X(gx, y), d_X(g^{-1}x, y)) > d(x, y)$$

then the action is diffuse.

Proof. Let $A \subset X$ be compact with at least two elements. Take a, b in A with $d(a, b) = \mathrm{diam(A)}$, then these are extremal in A. It suffices to check this for a. Given $g \in \Gamma$ not stabilizing a, then ga or $g^{-1}a$ is farther away from b, hence not in A. □

Note that this argument does not require nor that the action be isometric, neither that the function d_X on $X \times X$ be a distance. However this geometric statement is sufficient for all our concerns in this paper.

2.2. Related properties. Various properties of groups have been defined, which are closely related to diffuseness. We remind the reader of some of these properties and their mutual relations.

Let Γ be a group. We say that Γ is *locally indicable*, if every finitely generated non-trivial subgroup admits a non-trivial homomorphism into the group $\mathbb{Z}$. In

other words, every non-trivial finitely generated subgroup of Γ has a positive first rational Betti number.
Let $\prec$ be a total order on Γ. The order is called *left invariant*, if

$$x \prec y \implies gx \prec gy$$

for all x, y and g in Γ. We say that the order $\prec$ on Γ is *locally invariant* if for all $x, g \in \Gamma$ with $g \neq 1$ either $gx \prec x$ or $g^{-1}x \prec x$. Not all torsion-free groups admit orders with one of these properties. We say that Γ is *left-orderable* (resp. LIO) if there exists a left-invariant (resp. locally invariant) order on Γ. It is easily seen that an LIO group is diffuse. In fact, it was pointed out by Linnell and Witte Morris [46] that a group is LIO if and only if it is diffuse. One can see this as follows: If Γ is diffuse then every finite subset admits a locally invariant order (in an appropriate sense), and this yields a locally invariant order on Γ by a compactness argument.
The group Γ is said to have the *unique product property* (or to have unique products) if for every two finite non-empty subsets $A, B \subset \Gamma$ there is an element in the product $x \in A \cdot B$ which can be written uniquely in the form $x = ab$ with $a \in A$ and $b \in B$.
The following implications are well-known (for a complete account see [25]):

$$\text{locally indicable} \overset{(1)}{\Longrightarrow} \text{left-orderable} \overset{(2)}{\Longrightarrow} \text{diffuse} \overset{(3)}{\Longrightarrow} \text{unique products}$$

An example of Bergman [6] shows that (1) is in general not an equivalence, i.e. there are left-orderable groups which are not locally indicable (further examples are given by some of the hyperbolic three–manifolds studied in [16, Section 10] which have a left-orderable fundamental group with finite abelianization).
An explicit example showing that (2) is not an equivalence either is explained in the appendix written by Nathan Dunfield (see Theorem B.1). However, the reverse implication to (3), that is the relation between unique products and diffuseness, remains completely mysterious to us. We have no idea what the answer to the following question should be (even by restricting to groups in a smaller class, for example crystallographic, amenable, linear or hyperbolic groups).

Question 1. Does there exist a group which is not diffuse but has unique products?

It seems extremely hard to verify, for a given group, the unique product property without using any of the other three properties.

2.3. Some particular hyperbolic three–manifolds.

2.3.1. *A diffuse, non-orderable group.* In Appendix B Nathan Dunfield describes explicitly an example of an arithmetic Kleinian group which is diffuse but not left-orderable – this yields the following result (Theorem B.1).

Theorem 2.2 (Dunfield). *There exists a finitely presented (hyperbolic) group which is diffuse but not left-orderable.*

With Linnell and Witte-Morris' result this shows that there is a difference in these matters between amenable and hyperbolic groups. To verify that the group is diffuse one can use Bowditch's result or our Proposition 4.5.
Let us make a few comments on the origins of this example. The possibility to find such a group among this class of examples was proposed, unbeknownst to the authors, by A. Navas—see [25, 1.4.3]. Nathan Dunfield had previously computed a vast list of examples of closed hyperbolic three–manifolds whose fundamental group is not left-orderable (for some examples see [16]), using an algorithm described in the second paper. The example in Appendix B was not in this list, but was obtained by searching through the towers of finite covers of hyperbolic 3-manifolds studied in [18, §6].

2.3.2. *A non-diffuse lattice in* $\mathrm{PSL}_2(\mathbb{C})$. We also found an example of a compact hyperbolic 3-manifold with a non-diffuse fundamental group; in fact it is the hyperbolic three–manifold of smallest volume.

Theorem 2.3. *The fundamental group of the Weeks manifold is not diffuse.*

We verified this result by explicitly computing a ravel in the fundamental group of the Weeks manifold. We describe the algorithm and its implementations in Section A.1. In fact, given a group Γ and a finite subset A one can decide whether A contains a ravel by the following procedure: choose a random point $a \in A$; if it is extremal (which we check using a sub-algorithm based on the solution to the word problem in Γ) we iterate the algorithm on $A \setminus \{a\}$, otherwise we continue with another one. Once all the points of A have been tested, what remains is either empty or a ravel in Γ.

2.3.3. *Arithmetic Kleinian groups.* In a follow-up to this paper we will investigate the diffuseness properties of arithmetic Kleinian groups, in the hope of finding more examples of the above phenomena. Let us mention two results that will be proven there:

(i) Let $p > 2$ be a prime. There is a constant C_p such that if Γ is a torsion-free arithmetic group with invariant trace field F of degree p and discriminant $D_F > C_p$, then Γ is diffuse.
(ii) If Γ is a torsion-free Kleinian group derived from a quaternion algebra over an imaginary quadratic field F such that

$$D_F \neq -3, -4, -7, -8, -11, -15, -20, -24$$

then Γ is diffuse.

2.4. Groups which are not virtually diffuse. All groups considered in this article are residually finite and turn out to be virtually diffuse. Due to a lack of examples, we are curious about an answer to the following question.

Question 2. Is there a finitely generated (resp. finitely presented) group which is torsion-free, residually finite and not virtually diffuse?

The answer is positive without the finiteness hypotheses: given any non-diffuse, torsion-free, residually finite group Γ, then an infinite restricted direct product of factors isomorphic to Γ is residually finite and not virtually diffuse.

Furthermore, if we do not require the group to be residually finite, then one may take a restricted wreath product $\Gamma \wr U$ with some infinite group U. The group $\Gamma \wr U$ is not virtually diffuse and it is finitely generated if Γ and U are finitely generated (not finitely presented, however). Moreover, by a theorem of Gruenberg [32] such a wreath product (Γ non-abelian, U infinite) is not residually finite. Other examples of groups which are not virtually diffuse are the amenable simple groups constructed by K. Juschenko and N. Monod in [42]; these groups cannot be locally indicable, however they are neither residually finite nor finitely presented.

In the case of hyperbolic groups, this question is related to the residual properties of these groups – namely it is still not known if all hyperbolic groups are residually finite. A hyperbolic group which is not virtually diffuse would thus be, in light of the results of Delzant–Bowditch, not residually finite. It is unclear to the authors if this approach is feasible; for results in this direction see [31].

Finally, let us note that it would also be interesting to study the more restrictive class of linear groups instead of residually finite ones.

3. Fundamental groups of infra-solvmanifolds

3.1. Introduction.

3.1.1. *Infra-solvmanifolds.* In this section we discuss diffuse and non-diffuse fundamental groups of infra-solvmanifolds. The focus lies on crystallographic groups, however we shall begin the discussion in a more general setting. Let G be a connected, simply connected, solvable Lie group and let $\mathrm{Aut}(G)$ denote the group of continuous automorphisms of G. The affine group of G is the semidirect product $\mathrm{Aff}(G) = G \rtimes \mathrm{Aut}(G)$. A *lattice* $\Gamma \subset G$ is a discrete cocompact subgroup of G. An *infra-solvmanifold* (of type G) is a quotient manifold G/Λ where $\Lambda \subseteq \mathrm{Aff}(G)$ is a torsion-free subgroup of the affine group such that $\Lambda \cap G$ has finite index in Λ and is a lattice in G. If Λ is not diffuse, we say that G/Λ is a non-diffuse infra-solvmanifold.

The compact infra-solvmanifolds which come from a nilpotent Lie group G are characterised by the property that they are *almost flat*: that is, they admit Riemannian metrics with bounded diameter and arbitrarily small sectional curvatures (this is a theorem of M. Gromov, see [30], [15]). Those that come from abelian G are exactly those that are *flat*, i.e. they admit a Riemannian metric with vanishing sectional curvatures. We will study the latter in detail further in this section. We are not aware of any geometric characterization of general infra-solvmanifolds.

3.1.2. *Diffuse virtually polycyclic groups are strongly polycyclic.* Recall that a group Γ is (strongly) polycyclic if it admits a subnormal series with (infinite) cyclic factors. By a result of Mostow lattices in connected solvable Lie groups

are polycyclic (cf. Prop. 3.7 in [55]). Consequently, the fundamental group of an infra-solvmanifold is a virtually polycyclic group.
As virtually polycyclic groups are amenable, we can use the following striking result of Linnell and Witte Morris [46].

Theorem 3.1 (Linnell, Witte Morris). *An amenable group is diffuse if and only if it is locally indicable.*

We shall give a geometric proof of this theorem for the special case of virtually abelian groups in the next section. Here we confine ourselves to pointing out the following algebraic consequence.

Proposition 3.2. *A virtually polycyclic group Γ is diffuse if and only if Γ is strongly polycyclic. Consequently, the fundamental group of an infra-solvmanifold is diffuse exactly if it is strongly polycyclic.*

Proof. Clearly, a strongly polycyclic group is a virtually polycyclic group, in addition it is diffuse by Theorem 1.2 in [10].
Assume that Γ is diffuse and virtually polycyclic. We show that Γ is strongly polycyclic by induction on the Hirsch length $h(\Gamma)$. If $h(\Gamma) = 0$, then Γ is a finite group and as such it can only be diffuse if it is trivial.
Suppose $h(\Gamma) = n > 0$ and suppose that the claim holds for all groups of Hirsch length at most $n - 1$. By Theorem 3.1 the group Γ is locally indicable and (since Γ is finitely generated) we can find a surjective homomorphism $\phi\colon \Gamma \to \mathbb{Z}$. Observe that $h(\Gamma) = h(\ker(\phi)) + 1$. The kernel $\ker(\phi)$ is diffuse and virtually polycyclic, and we deduce from the induction hypothesis, that $\ker(\phi)$ (and so Γ) is strongly polycyclic. $\square$

In the next three sections we focus on crystallographic groups. After the discussion of a geometric proof of Theorem 3.1 in the crystallographic setting (3.2), we will analyse the influence of the structure of the holonomy group for the existence of ravels (3.3). We also give a list of all non-diffuse crystallographic groups in dimension up to four (3.4). Finally, we discuss a family of non-diffuse infra-solvmanifolds in 3.5 which are not flat manifolds.

3.2. Geometric construction of ravels in virtually abelian groups.

The equivalence of local indicability and diffuseness for amenable groups which was established by Linnell and Witte Morris [46] is a powerful result. Accordingly a virtually polycyclic group with vanishing first rational Betti number contains a ravel. However, their proof does not explain a construction of ravels based on the vanishing Betti number. They stress that this does not seem to be obvious even for virtually abelian groups. The purpose of this section is to give a geometric and elementary proof of this theorem, for the special case of virtually abelian groups, which is based on an explicit construction of ravels.

Theorem 3.3. *A virtually abelian group is diffuse exactly if it is locally indicable.*

As discussed in Section 2.2 local indicability implies diffuseness. It suffices to prove the converse. Let Γ_0 be a virtually abelian group and assume that it is not locally indicable. We can find a finitely generated subgroup $\Gamma \subset \Gamma_0$ with vanishing first rational Betti number. If Γ contains torsion, it is not diffuse. Thus we assume that Γ is torsion-free. Since a finitely generated torsion-free virtually abelian group is crystallographic, the theorem follows from the next lemma.

Lemma 3.4. *Let Γ be a crystallographic group acting on a euclidean space E. If $b_1(\Gamma) = 0$, then for all $e \in E$ and all sufficiently large $r > 0$ the set*

$$B(r, e) = \{ \gamma \in \Gamma \mid \|\gamma e - e\| \leq r \}$$

is a ravel.

Proof. We can assume $e = 0 \in E$. Let Γ be a non-trivial crystallographic group with vanishing first Betti number and let $\pi : \Gamma \to G$ be the projection onto the holonomy group at 0. The translation subgroup is denoted by T and we fix some $r_0 > 0$ so that for every $u \in E$ there is $t \in T$ satisfying $\|u - t\| \leq r_0$. The first Betti number $b_1(\Gamma)$ is exactly the dimension of the space E^G of G-fixed vectors. Thus $b_1(\Gamma) = 0$ means that G acts without non-trivial fixed points on E. Since every non-zero vector is moved by G, there is a real number $\delta < 1$ such that for all $u \in E$ there is $g \in G$ such that

$$\|gu + u\| \leq 2\delta\|u\|. \tag{1}$$

For $r > 0$ let B_r denote the closed ball of radius r around 0. Fix $u \in B_r$; we shall find $\gamma \in \Gamma$ such that $\|\gamma u\| \leq r$ and $\|\gamma^{-1}u\| \leq r$ provided r is sufficiently large. We pick $g \in G$ as in (1) and we choose some $\gamma_0 \in \Gamma$ with $\pi(\gamma_0) = g$. Define $w_0 = \gamma_0(0)$. We observe that for every two vectors $v_1, v_2 \in E$ with distance d, there is $x \in w_0 + T$ with

$$\max_{i=1,2}(\|v_i - x\|) \leq r_0 + \frac{d}{2}.$$

Indeed, the ball of radius r_0 around the midpoint of the line between v_1 and v_2 contains an element $x \in w_0 + T$. Apply this to the vectors $v_1 = u$ and $v_2 = -gu$ to find some $x = w_0 + t$. By construction we get $d \leq 2\delta r$. Finally we define $\gamma = t \circ \gamma_0$ to deduce the inequalities

$$\|\gamma u\| = \|gu + x\| = \| - gu - x\| \leq r_0 + \delta r$$

and

$$\|\gamma^{-1}u\| = \|g^{-1}u - g^{-1}x\| = \|u - x\| \leq r_0 + \delta r.$$

As $\delta < 1$ the right hand side is less than r for all sufficiently large r. $\square$

3.3. Diffuseness and the holonomy of crystallographic groups. We take a closer look at the non-diffuse crystallographic groups and their holonomy groups. It will turn out that for a given crystallographic group one can often decide from the holonomy group whether or not the group is diffuse. In the following a *Bieberbach group* is a non-trivial torsion-free crystallographic group. Let Γ be a Bieberbach group, it has a finite index normal maximal abelian subgroup $T \subset \Gamma$. Recall that the finite quotient $G = \Gamma/T$ is called the holonomy group of Γ. Since every finite group is the holonomy group of some Bieberbach group (by a result due to Auslander-Kuranishi [4]), this naturally divides the finite groups into three classes.

Definition 1. A finite group G is *holonomy diffuse* if every Bieberbach group Γ with holonomy group G is diffuse. It is *holonomy anti-diffuse* if every Bieberbach group Γ with holonomy group G is non-diffuse. Otherwise we say that G is *holonomy mixed*.

For example, the finite group $(\mathbb{Z}/2\mathbb{Z})^2$ is holonomy mixed. In fact, the Promislow group Δ_P (also known as Hantzche-Wendt group or Passman group) is a non-diffuse [10] Bieberbach group with holonomy group $(\mathbb{Z}/2\mathbb{Z})^2$ – thus $(\mathbb{Z}/2\mathbb{Z})^2$ is not holonomy diffuse. On the other hand it is easy to construct diffuse groups with holonomy group $(\mathbb{Z}/2\mathbb{Z})^2$ (cf. Lemma 3.9 below).
In this section we prove the following algebraic characterisation of these three classes of finite groups.

Theorem 3.5. *A finite group G is*

(i) holonomy anti-diffuse if and only if it is not solvable.
(ii) holonomy diffuse exactly if every Sylow subgroup is cyclic.
(iii) holonomy mixed if and only if it is solvable and has a non-cyclic Sylow subgroup.

The proof of this theorem will be given as a sequence of lemmata below. A finite group G with cyclic Sylow subgroups is meta-cyclic (Thm. 9.4.3 in [34]). In particular, such a group G is solvable and hence it suffices to prove the assertions (i) and (ii). One direction of (i) is easy. By Proposition 3.2 a diffuse Bieberbach group is solvable and thus cannot have a finite non-solvable quotient, i.e. a non-solvable group is holonomy anti-diffuse. For (i) it remains to verify that every finite solvable group is the holonomy of some diffuse Bieberbach group; this will be done in Lemma 3.9.
In order to prove (ii), we shall use a terminology introduced by Hiller-Sah [38].

Definition 2. A finite group G is *primitive* if it is the holonomy group of a Bieberbach group with finite abelianization.

Statement (ii) of the theorem will follow from the next lemma.

Lemma 3.6. *Let G be a finite group. The following statements are equivalent.*

(a) G is not holonomy diffuse.
(b) G has a non-cyclic Sylow subgroup.
(c) G contains a normal primitive subgroup.

We frequently use the following notion: A cohomology class $\alpha \in H^2(G, A)$ (for some finite group G and some G-module A) is called *special* if it corresponds to a torsion-free extension of G by A (cf. [38]). Equivalently, if A is free abelian, the restriction of α to any cyclic subgroup of G is non-zero.

Proof. Hiller-Sah [38] obtained an algebraic characterisation of primitive groups. They showed that a finite group is primitive exactly if it does not contain a cyclic Sylow p-subgroup which admits a normal complement (see also [21] for a different criterion).
(a) $\implies$ (b): Assume G is not holonomy diffuse and take a non-diffuse Bieberbach group Γ with holonomy group G. As Γ is not locally indicable we find a non-trivial subgroup $\Gamma_0 \leq \Gamma$ with $b_1(\Gamma_0) = 0$. The holonomy group G_0 of Γ_0 is primitive. Let p be the smallest prime divisor of $|G_0|$. The Sylow p-subgroups of G_0 are not cyclic, since otherwise they would admit a normal complement (by a result of Burnside [14]). Let $\pi : \Gamma \to G$ be the projection. The image $\pi(\Gamma_0)$ has G_0 as a quotient and hence $\pi(\Gamma_0)$ also has non-cyclic Sylow p-subgroups. As every p-group is contained in a Sylow p-subgroup, we deduce that the Sylow p-subgroups of G are not cyclic.

(b) $\implies$ (c): Let p be a prime such that the Sylow p-subgroups of G are not cyclic. Consider the subgroup H of G generated by all p-Sylow subgroups. The group H is normal in G and we claim that it is primitive. The Sylow p-subgroups of H are precisely those of G and they are not cyclic. Let p' be a prime divisor of $|H|$ different from p. Suppose there is a (cyclic) Sylow p'-subgroup Q in H which admits a normal complement N. As H/N is a p'-group, the Sylow p-subgroups of H lie in N. By construction H is generated by its Sylow p-subgroups and so $N = H$. This contradicts the existence of such a Sylow p'-subgroup.

(c) $\implies$ (a): Assume now that G contains a normal subgroup $N \trianglelefteq G$ which is primitive. We show that G is not holonomy diffuse. Since N is primitive, there exists Bieberbach group Λ with holonomy group N and with $b_1(\Lambda, \mathbb{Q}) = 0$. Let A be the translation subgroup of Λ and let $\alpha \in H^2(N, A)$ be the special class corresponding to the extension Λ. The vanishing Betti number $b_1(\Lambda, \mathbb{Q}) = 0$ is equivalent to $A^N = \{0\}$.
Consider the induced $\mathbb{Z}[G]$-module $B := \mathrm{ind}_N^G(A)$. Let T be a transversal of N in G containing 1_G. If we restrict the action on B to N we obtain

$$B_{|N} = \bigoplus_{g \in T} A(g)$$

where $A(g)$ is the N-module obtained from A by twisting with the action with g, i.e. $h \in N$ acts by $g^{-1}hg$ on A. In particular, $B^N = \{0\}$ and $A = A(1_G)$ is a direct summand of $B_{|N}$.
Observe that every class in $H^2(N, B)$ which projects to $\alpha \in H^2(N, A)$ is special and defines thus a Bieberbach group with finite abelianization. Shapiro's isomorphism $\mathrm{sh}^2 \colon H^2(G, B) \to H^2(N, A)$ is the composition of the restriction res_G^N and the projection $H^2(N, B) \to H^2(N, A)$. We deduce that there is a class $\gamma \in H^2(G, B)$ which maps to some special class $\beta \in H^2(N, B)$ (which projects onto $\alpha \in H^2(N, A)$). Let Λ' be the Bieberbach group (with $b_1(\Lambda') = 0$) corresponding to β. The group corresponding to γ might not be torsion-free, so we need to vary γ so that it becomes a special class.

Let $\mathfrak{H}$ be the collection of all cyclic prime order subgroups C of G which intersect N trivially. For each $C \in \mathfrak{H}$ we define

$$M_C := \mathrm{ind}_C^G(\mathbb{Z})$$

where C acts trivially on $\mathbb{Z}$. The group N acts freely on $C\backslash G$, since $C \cap N = \{1_G\}$. Therefore $(M_C)_{|N}$ is a free $\mathbb{Z}[N]$-module. We define the $\mathbb{Z}[G]$-module

$$M = B \oplus \bigoplus_{C \in \mathfrak{H}} M_C.$$

Using Shapiro's Lemma we find classes $\alpha_C \in H^2(G, M_C)$ which restrict to non-trivial classes in $H^2(C, M_C)$. Consider the cohomology class $\delta := \gamma \oplus \bigoplus_{C \in \mathfrak{H}} \alpha_C \in H^2(G, M)$.
The class δ is special, as can be seen as follows. For every $C \in \mathfrak{H}$ this follows from the fact that α_C restricts non-trivially to C. For the cyclic subgroups $C \leq N$ this holds since the restriction of γ to N is special. Consequently δ defines a Bieberbach group Γ with holonomy group G.
Finally, we claim that $\mathrm{res}_G^N(\delta) = i_*(\mathrm{res}_G^N(\gamma))$ where $i : B \to M$ is the inclusion map. Indeed, $H^2(N, M_C) = 0$ since M_C is a free $\mathbb{Z}[N]$-module. Since $\mathrm{res}_G^N(\gamma) = \beta$ we conclude that Γ contains the group Λ' as a subgroup and thus Γ is not locally indicable. $\square$

We are left with constructing diffuse Bieberbach groups for a given solvable holonomy group. We start with a simple lemma concerning fibre products of groups. For $0 \leq i \leq n$ let Γ_i be a group with a surjective homomorphism ψ_i onto some fixed group G. The fibre product $\times_G \Gamma_i$ is defined as a subgroup of the direct product $\prod_i \Gamma_i$ by

$$\times_G \Gamma_i := \{\, (\gamma_i)_i \in \prod_{i=0}^{n} \Gamma_i \mid \psi_i(\gamma_i) = \psi_0(\gamma_0) \ \text{ for all } i \,\}.$$

In this setting we observe the following

Lemma 3.7. *If Γ_0 is diffuse and $\ker \psi_i \subset \Gamma_i$ is diffuse for all $i \in \{1, \ldots, n\}$, then $\times_G \Gamma_i$ is diffuse.*

Proof. There is a short exact sequence

$$1 \longrightarrow \prod_{i=1}^{n} \ker \psi_i \xrightarrow{\ j\ } \times_G \Gamma_i \longrightarrow \Gamma_0 \longrightarrow 1$$

so the claim follows from Theorem 1.2 in [10]. $\square$

Lemma 3.8. *Let G be a finite group and let $M_1, \ldots, M_n$ be free $\mathbb{Z}$-modules with G-action. Let $\alpha_i \in H^2(G, M_i)$ be classes. If one of these classes defines a diffuse extension group of G, then the sum of the α_i in $H^2(G, M_1 \oplus \cdots \oplus M_n)$ defines a diffuse extension of G.*

Proof. Taking the sum of classes corresponds to the formation of fibre products of the associated extensions, so the claim follows from Lemma 3.7. $\square$

Lemma 3.9. *Every finite solvable group is the holonomy group of a diffuse Bieberbach group.*

Proof. We begin by constructing diffuse Bieberbach groups with given abelian holonomy group. Let A be an abelian group and let Γ_1 be a Bieberbach group with holonomy group A and projection $\psi_1\colon \Gamma_1 \to A$. Write A as a quotient of a free abelian group $\Gamma_0 = \mathbb{Z}^k$ of finite rank with projection $\psi_0\colon \mathbb{Z}^k \to A$. By Lemma 3.7 the fibred product $\Gamma_0 \times_A \Gamma_1$ is a diffuse Bieberbach group with holonomy group A (the kernel of ψ_1 is free abelian).
Assume now that G is solvable. We construct a diffuse Bieberbach group Γ with holonomy group G. We will proceed by induction on the derived length of G. The basis for the induction is given by the construction for abelian groups above. Let G' be the derived group of G. By induction hypothesis there is a faithful G'-module M and a "diffuse" class $\alpha \in H^2(G', M)$. Consider the induced module $B = \mathrm{ind}_{G'}^{G}(M)$. The restriction of B to G' decomposes into a direct sum

$$B_{|G'} \cong M \oplus X.$$

There is a class $\beta \in H^2(G, B)$ which maps to α under Shapiro's isomorphism $\mathrm{sh}^2\colon H^2(G,B) \to H^2(G', M)$. Due to this the restriction $\mathrm{res}_G^{G'}(\beta)$ decomposes as $\alpha \oplus x \in H^2(G', M) \oplus H^2(G', X)$. By Lemma 3.8 the class $\mathrm{res}_G^{G'}(\beta)$ is diffuse. Let Γ_1 be the extension of G which corresponds to the class β. By what we have seen, the subgroup $\Lambda_1 = \ker(\Gamma_1 \to G/G')$ is diffuse. Finally, we write the finite abelian group G/G' as a quotient of a free abelian group $\Gamma_0 = \mathbb{Z}^k$. By Lemma 3.7 the fibre product $\Gamma_0 \times_{G/G'} \Gamma_1$ is diffuse. In fact, it is a Bieberbach group with holonomy group G. □

3.4. Non-diffuse Bieberbach groups in small dimensions. In this section we briefly describe the classification of all Bieberbach groups in dimension $d \leq 4$ which are not diffuse. The complete classification of crystallographic groups in these dimensions is given in [13] and we refer to them according to their system of enumeration.
In dimensions 2 and 3 the classification is very easy. In dimension $d = 2$ there are two Bieberbach groups and both of them are diffuse. In dimension $d = 3$ there are exactly 10 Bieberbach groups. The only group among those with vanishing first rational Betti number is the Promislow (or Hantzsche-Wendt) group Δ_P (which is called 3/1/1/04 in [13]).
Now we consider the case $d = 4$, in this case there are 74 Bieberbach groups. As a consequence of the considerations for dimensions 2 and 3, a Bieberbach group Γ of dimension $d = 4$ is not diffuse if and only if it has vanishing Betti number or contains the Promislow group Δ_P. Vanishing Betti number is something that can be detected easily from the classification. So how can one detect the existence of a subgroup isomorphic to Δ_P? The answer is given in the following lemma.

Lemma 3.10. *Let Γ be a Bieberbach group acting on $E = \mathbb{R}^4$ and assume that $b_1(\Gamma) > 0$. Let $\pi : \Gamma \to G$ be the projection onto the holonomy group. Then Γ is not diffuse if and only if it contains elements $g, h \in \Gamma$ such that*

(i) $S := \langle \pi(g), \pi(h) \rangle \cong (\mathbb{Z}/2\mathbb{Z})^2$,
(ii) $\dim E^S = 1$ *and*
(iii) if $E = E^S \oplus V$ *as S-module, then $g \cdot 0$ and $h \cdot 0$ lie in V.*

Proof. Since $b_1(\Gamma) > 0$, the group Γ is not diffuse exactly if it contains Δ_P as a subgroup.
Assume Γ contains Δ_P and let $\Lambda = \ker(\pi)$ be the translation subgroup of Γ (considered as a lattice in E). We claim that the holonomy group of Δ_P embeds into G via π. We show that $L := \Delta_P \cap \Lambda$ is the maximal abelian finite index subgroup of Δ_P. The lattice L spans a three-dimensional subspace $V \subseteq E$ on which $\Delta_P/L \cong \pi(\Delta_P)$ acts without fixed points. Since $b_1(\Gamma) > 0$ the group $S = \pi(\Delta_P)$ has a one-dimensional fixed point space E^S which is a complement of V in E. Suppose L_1 is an abelian subgroup of Δ_P which contains L. Then L_1/L acts trivially on E and (as G acts faithfully on E) we conclude $L_1 = L$. Take g and h in Γ such $\pi(g)$ and $\pi(h)$ generate S, clearly $g \cdot 0, h \cdot 0 \in V$. The group S acts without non-trivial fixed points on V and $E = E^S \oplus V$ is a decomposition as S-module.
Conversely, if we can find $g, h \in \Gamma$ as above, then they generate a Bieberbach group of smaller dimension and with vanishing first Betti number. Hence they generate a group isomorphic to Δ_P. □

Using this lemma and the results of the previous section one can decide for each of the 74 Bieberbach groups whether they are diffuse or not. It turns out there are 16 non-diffuse groups in dimension 4, namely (cf. [13]):

04/03/01/006, 05/01/02/009, 05/01/04/006, 05/01/07/004,
06/01/01/049, 06/01/01/092, 06/02/01/027, 06/02/01/050,
12/03/04/006, 12/03/10/005, 12/04/03/011, 13/04/01/023,
13/04/04/011, 24/01/02/004, 24/01/04/004, 25/01/01/010.

The elementary abelian groups $(\mathbb{Z}/2\mathbb{Z})^2$, $(\mathbb{Z}/2\mathbb{Z})^3$, the dihedral group D_8, the alternating group A_4 and the direct product group $A_4 \times \mathbb{Z}/2\mathbb{Z}$ occur as holonomy groups. Among these groups only four groups have vanishing first Betti number (these are 04/03/01/006, 06/02/01/027, 06/02/01/050 and 12/04/03/011). However, one can check that these groups contain the Promislow group as well. In a sense the Promislow group is the only reason for Bieberbach groups in dimension 4 to be non-diffuse (thus non of these groups has the unique product property). This leads to the following question: What is the smallest dimension d_0 of a non-diffuse Bieberbach group which does not contain Δ_P? Clearly, such a group has vanishing first Betti number. Note that there is a group with vanishing first Betti number and holonomy $(\mathbb{Z}/3\mathbb{Z})^2$ in dimension 8 (see [38]); thus $5 \leq d_0 \leq 8$. The so-called generalized Hantzsche-Wendt groups are higher dimensional analogs of Δ_P (cf. [63, 58]). However, any such group Γ with $b_1(\Gamma) = 0$ contains the Promislow group (see Prop. 8.2 in [58]).

3.5. A family of non-diffuse infra-solvmanifolds. Many geometric questions are not answered by the simple algebraic observation in Proposition 3.2. For instance, given a simply connected solvable Lie group G, is there an infra-solvmanifold of type G with non-diffuse fundamental group? To our knowledge there is no criterion which decides whether a solvable Lie group G admits a lattice at all. Hence we do not expect a simple answer for the above question. We briefly discuss an infinite family of simply connected solvable groups where every infra-solvmanifold is commensurable to a non-diffuse one. Let $\rho_1, \dots, \rho_n$ be $n \geq 1$ *distinct* real numbers with $\rho_i > 1$ for all $i = 1, \dots, n$. We define the Lie group

$$G := \mathbb{R}^{2n} \rtimes \mathbb{R}$$

where $s \in \mathbb{R}$ acts by the diagonal matrix $\beta(s) := \mathrm{diag}(\rho_1^s, \dots, \rho_n^s, \rho_1^{-s}, \dots, \rho_n^{-s})$ on $\mathbb{R}^{2n}$. The group G is a simply connected solvable Lie group. The isomorphism class of G depends only one the line spanned by $(\log \rho_1, \dots, \log \rho_n)$ in $\mathbb{R}^n$. For $n = 1$ the group G is the three dimensional solvable group **Sol**, which will be reconsidered in Section 5.

Proposition 3.11. *In the above setting the following holds.*

(a) *The Lie group G has a lattice if and only if there is $t_0 > 0$ such that the polynomial $f(X) := \prod_{i=1}^n (1 - (\rho_i^{t_0} + \rho_i^{-t_0})X + X^2)$ has integral coefficients.*

(b) *If G admits a lattice, then every infra-solvmanifold of type G is commensurable to a non-diffuse one.*

Before we prove the proposition, we describe the group of automorphisms of G. Let $\sigma \in \mathrm{Aut}(G)$, then $\sigma(x, t) = (Wx + f(t), \lambda t)$ for some $\lambda \in \mathbb{R}^\times$, $W \in \mathrm{GL}_{2n}(\mathbb{R})$ and $f \in Z^1(\mathbb{R}, \mathbb{R}^{2n})$ a smooth cocycle for the action of $s \in \mathbb{R}$ on $\mathbb{R}^{2n}$ via $\beta(\lambda \cdot s)$. Using that $H^1(\mathbb{R}, \mathbb{R}^{2n}) = 0$ we can compose σ with an inner automorphism of G (given by an element in $[G, G]$) such that $f(t) = 0$. Observe that the following equality has to hold

$$\beta(\lambda t)W = W\beta(t)$$

for all $t \in \mathbb{R}$. As a consequence λ is 1 or -1. In the former case W is diagonal, in the latter case W is a product of a diagonal matrix and

$$W_0 = \begin{pmatrix} 0 & 1_n \\ 1_n & 0 \end{pmatrix}.$$

Let D_+ denote the group generated by diagonal matrices in $\mathrm{GL}_{2n}(\mathbb{R})$ and W_0, then $\mathrm{Aut}(G) \cong \mathbb{R}^{2n} \rtimes D_+$.

Proof of Proposition 3.11. Ad (a): Note that $N := \mathbb{R}^{2n} = [G, G]$ is the maximal connected normal nilpotent subgroup of G. Suppose that G contains a lattice Γ. Then $\Gamma_0 := \Gamma \cap N$ is a lattice in N (cf. Cor. 3.5 in [55]) and Γ/Γ_0 is a lattice in $G/N \cong \mathbb{R}$. Let $t_0 \in \mathbb{R}$ so that we can identify Γ/Γ_0 with $\mathbb{Z}t_0$ in $\mathbb{R}$. Take a basis of Γ_0, with respect to this basis $\beta(t_0)$ is a matrix in $\mathrm{SL}_{2n}(\mathbb{Z})$. The polynomial f is the charcteristic polynomial of $\beta(t_0)$ and the claim follows.

Conversely, let $t_0 > 0$ with $f \in \mathbb{Z}[X]$ as above. Take any matrix $A \in \mathrm{SL}_{2n}(\mathbb{Z})$ with characteristic polynomial f, e.g. if $f(X) = X^{2n} + a_{2n-1}X^{2n-1} + \cdots + a_1 X + a_0$ then the matrix A with ones above the diagonal and last row $(-a_0, -a_1, \ldots, -a_{2n-1})$ has suitable characteristic polynomial.
Since by assumption all the ρ_i are distinct and real, we find $P \in \mathrm{GL}_{2n}(\mathbb{R})$ with $PAP^{-1} = \beta(t_0)$. Now define $\Gamma_0 := P\mathbb{Z}^{2n}$ and we obtain a lattice $\Gamma := \Gamma_0 \rtimes (\mathbb{Z}t_0)$ in G.

Ad (b): Let $\Lambda \subset \mathrm{Aff}(G)$ be the fundamental group of an infra-solvmanifold. Define $\Gamma := G \cap \Lambda$ and $\Gamma_0 := \Gamma \cap N$ where $N = \mathbb{R}^{2n}$ is the maximal normal nilpotent subgroup. The first Betti number of Λ is $b_1(\Lambda) = \dim_{\mathbb{R}}(G/N)^{\Lambda/\Gamma}$. The quotient Γ/Γ_0 is a lattice in $\mathbb{R}$, so is of the form $\mathbb{Z}t_0$ for some $t_0 > 0$.
Take any basis of the lattice $\Gamma_0 \subseteq \mathbb{R}^{2n}$. We shall consider coordinates on $\mathbb{R}^{2n}$ with respect to this basis from now on. In particular, $\beta(t_0)$ is given by an integral matrix $A \in \mathrm{SL}_{2n}(\mathbb{Z})$ and further Γ is isomorphic to the strongly polycyclic group $\mathbb{Z}^{2n} \rtimes \mathbb{Z}$ where $\mathbb{Z}$ acts via A. Let $F/\mathbb{Q}$ be a finite totally real Galois extension which splits the characteristic polynomial of A, so the Galois group permutes the eigenvalues of A. Moreover, the Galois group acts on $\Gamma_0 \otimes_{\mathbb{Z}} F$ so that we can find a set of eigenvectors which are permuted accordingly. Let $B \in \mathrm{GL}_{2n}(F)$ be the matrix whose columns are the chosen eigenvectors, then $B^{-1}AB = \beta(t_0)$ and for all $\sigma \in \mathrm{Gal}(F/\mathbb{Q})$ we have $\sigma(B) = BP_\sigma$ for a permutation matrix $P_\sigma \in \mathrm{GL}_{2n}(\mathbb{Z})$. It is easily seen that P_σ commutes with W_0, and hence $W = BW_0B^{-1}$ is stable under the Galois group, this means $W \in \mathrm{GL}_{2n}(\mathbb{Q})$.
Since W is of order two, we can find a sublattice $L \subset \Gamma_0$ which admits a basis of eigenvectors of W. Pick one of these basis vectors, say v, with eigenvalue one, find $q \in \mathbb{Z} \setminus \{0\}$ with $q\Gamma_0 \subset L$ and take a positive integer r so that

$$A^r \equiv 1 \bmod 4q.$$

This way we find a finite index subgroup $\Gamma' := L \rtimes r\mathbb{Z}$ of Γ which is stable under the automorphism τ defined by $(x, t) \mapsto (Wx, -t)$. Since we want to construct a torsion-free group we cannot add τ into the group. Instead we take the group Λ' generated by $(\frac{1}{2}v, 0)\tau$ and Γ' in the affine group $\mathrm{Aff}(G)$. A short calculation shows that Λ' is torsion-free and hence Λ' is the fundamental group of an infra-solvmanifold of type G which is commensurable with Λ. By construction the first Betti number $b_1(\Lambda') = \dim_{\mathbb{R}}(G/N)^{\Lambda'/\Gamma'}$ vanishes and so Λ' is not diffuse by Theorem 3.1. □

4. Fundamental groups of hyperbolic manifolds

In this section we prove Theorems B and C from the introduction. We give a short overview of rank one symmetric spaces before studying first their unipotent and then their axial isometries in view of applying Lemma 2.1. Then we review some well-known properties of geometrically finite groups of isometries before proving a more general result (Theorem 4.8) and showing how it implies

Theorems B and C. We also study the action on the boundary, resulting in Theorem 4.11, which will be used in the next section.

4.1. Hyperbolic spaces.

4.1.1. *Isometries.* We recall some terminology about isometries of Hadamard manifolds: if $g \in \mathrm{Isom}^+(X)$ where X is a complete simply connected manifold with non-positive curvature then g is said to be

- Hyperbolic (or axial) if $\min(g) = \inf_{x \in X} d_X(x, gx) > 0$;
- Parabolic if it fixes exactly one point in the visual boundary ∂X, equivalently $\min(g) = 0$ and g has no fixed point inside X.

We will be interested here in the case where $X = G/K$ is a symmetric space associated to a simple Lie group G *of real rank one*. An element $g \in G$ then acts on X as an hyperbolic isometry if and only if it is semisimple and has an eigenvalue of absolute value > 1 in the adjoint representation. Parabolic isometries of X are algebraically characterised as corresponding to the non-semisimple elements of G; their eigenvalues are necessarily of absolute value one. If they are all equal to one then the element of G is said to be unipotent, as well as the corresponding isometry of X.

4.1.2. *Projective model.* Here we describe models for the hyperbolic spaces $\mathbb{H}^n_A$ for $A = \mathbb{R}, \mathbb{C}, \mathbb{H}$ (the symmetric spaces associated to the Lie groups $\mathrm{SO}(n,1)$, $\mathrm{SU}(n,1)$ and $\mathrm{Sp}(n,1)$ respectively) which we will use later for computations. We will denote by $z \mapsto \overline{z}$ the involution on A fixing $\mathbb{R}$, and define as usual the reduced norm and trace of A by

$$|z|_{A/\mathbb{R}} = \overline{z}z = z\overline{z}, \quad \mathrm{tr}_{A/\mathbb{R}}(z) = z + \overline{z}$$

We let $V = A^{n,1}$, by which we mean that V is the right A-vector space A^{n+1} endowed with the sesquilinear inner product given by[2]

$$\langle v, v' \rangle = \overline{v'_{n+1}} v_1 + \sum_{i=2}^{n} \overline{v'_i} v_i + \overline{v'_1} v_{n+1}.$$

The (special if $A = \mathbb{R}$ or $\mathbb{C}$) isometry group G of V is then isomorphic to $\mathrm{SO}(n,1)$, $\mathrm{SU}(n,1)$ or $\mathrm{Sp}(n,1)$. Let:

$$V_- = \{v \in V \mid \langle v, v \rangle < 0\} = \left\{ v \in V \mid \mathrm{tr}_{A/\mathbb{R}}(\overline{v_1} v_{n+1}) < -\sum_{i=2}^{n} |v_i|_{A/\mathbb{R}} \right\}$$

then the image $X = \mathbb{P}V_-$ of V_- in the A-projective space $\mathbb{P}V$ of V can be endowed with a distance function d_X given by:

$$\cosh\left(\frac{d_X([v],[v'])}{2}\right)^2 = \frac{|\langle v, v' \rangle|_{A/\mathbb{R}}}{\langle v, v \rangle \langle v', v' \rangle}. \tag{2}$$

This distance is G-invariant, and the stabilizer in G of a point in V_- is a maximal compact subgroup of G. Hence the space X is a model for the symmetric

[2] We use the model of [43] rather than that of [53].

space G/K (where $K = \mathrm{SO}(n), \mathrm{SU}(n)$ or $\mathrm{Sp}(n)$ according to whether $A = \mathbb{R}, \mathbb{C}$ or $\mathbb{H}$).

The following lemma will be of use later.

Lemma 4.1. *If $v, v' \in V_-$ then $\mathrm{tr}_{A/\mathbb{R}}(\overline{v_{n+1}} v'_{n+1} \langle v, v' \rangle) < 0$.*

Proof. Since $\mathrm{tr}_{A/\mathbb{R}}(\overline{v_{n+1}} v'_{n+1} \langle v, v' \rangle)$ does not change sign when we multiply v or v' by a element of A from the right, we may suppose that $v_{n+1} = v'_{n+1} = 1$. In this case we have:

$$\mathrm{tr}_{A/\mathbb{R}}(\overline{v_{n+1}} v'_{n+1} \langle v, v' \rangle) = \mathrm{tr}_{A/\mathbb{R}}(v_1) + \mathrm{tr}_{A/\mathbb{R}}(v'_1) + \mathrm{tr}_{A/\mathbb{R}} \left(\sum_{i=2}^{n} \overline{v'_i} v_i \right).$$

Now we have

$$\mathrm{tr}_{A/\mathbb{R}} \left(\sum_{i=2}^{n} \overline{v'_i} v_i \right) \leq 2 \sqrt{ \left(\sum_{i=2}^{n} |v_i|_{\mathrm{A}/\mathbb{R}} \right) \cdot \left(\sum_{i=2}^{n} |v'_i|_{\mathrm{A}/\mathbb{R}} \right) }$$

by Cauchy-Schwarz, and since $v, v' \in V_-$ we get

$$\begin{aligned} \mathrm{tr}_{A/\mathbb{R}}(\overline{v_{n+1}} v'_{n+1} \langle v, v' \rangle) &< \mathrm{tr}_{A/\mathbb{R}} \left(\sum_{i=2}^{n} \overline{v'_i} v_i \right) - \sum_{i=2}^{n} |v_i|_{\mathrm{A}/\mathbb{R}} - \sum_{i=2}^{n} |v'_i|_{\mathrm{A}/\mathbb{R}} \\ &\leq - \left(\sqrt{\sum_{i=2}^{n} |v_i|_{\mathrm{A}/\mathbb{R}}} - \sqrt{\sum_{i=2}^{n} |v'_i|_{\mathrm{A}/\mathbb{R}}} \right)^2 \leq 0. \end{aligned}$$

□

4.2. Unipotent isometries and distance functions. In this subsection we prove the following proposition, which is the main ingredient we use in extending the results of [10] from cocompact subgroups to general lattices.

Proposition 4.2. *Let A be one of $\mathbb{R}, \mathbb{C}$ or $\mathbb{H}$ and let $\eta \neq 1$ be a unipotent isometry of $X = \mathbb{H}^n_A$ and $a, x \in \mathbb{H}^n_A$. Then*

$$\max\big(d(a, \eta x), d(a, \eta^{-1} x)\big) > d(a, x).$$

Proof. We say that a function $h : \mathbb{Z} \to \mathbb{R}$ is strictly convex if h is the restriction to $\mathbb{Z}$ of a strictly convex function on $\mathbb{R}$ (equivalently all points on the graph of h are extremal in their convex hull and h has a finite lower bound). We will use the following criterion, similar to Lemma 6.1 in [53].

Lemma 4.3. *Let X be a metric space, $x \in X$ and let $\phi \in \mathrm{Isom}(X)$. Suppose that there exists an increasing function $f : [0, +\infty[\to \mathbb{R}$ such that for any $y \in X$ the function $h_y : k \mapsto f(d_X(y, \phi^k x))$ is strictly convex. Let*

$$B_k = \{ y \in X : \ d_X(y, \phi^k x) \leq d_X(y, x) \}.$$

Then we have $B_1 \cap B_{-1} = \emptyset$.

Proof. Suppose there is a $y \in X$ such that
$$d_X(y, \phi x), d_X(y, \phi^{-1}x) \le d_X(y, x).$$
Since f is increasing this means that $h_y(1), h_y(-1) \le h_y(0)$: but this is impossible since h_y is strictly convex. □

Applying it to $\phi = \eta$, we see that it suffices to prove that for any $z, w \in X$ the function
$$f : t \in \mathbb{R} \mapsto \cosh\left(\frac{d_X(z, \eta^t w)}{2}\right)^2$$
is strictly convex on $\mathbb{R}$, i.e. $f'' > 0$. Of course we need only to prove that $f''(0) > 0$ since z, w are arbitrary. By the formula (2) for arc length in hyperbolic spaces it suffices to prove this for the function
$$h : t \mapsto |\langle v, \eta^t v'\rangle|_{A/\mathbb{R}}$$
for any two $v, v' \in A^{n,1}$ (which we normalize so that their last coordinate equals 1). Now we have:
$$\begin{aligned}\frac{d^2h}{dt^2} &= \frac{d}{dt}\left(\mathrm{tr}_{A/\mathbb{R}}\left(\overline{\langle v, \eta^t v'\rangle}\frac{d}{dt}\langle v, \eta^t v'\rangle\right)\right)\\ &= 2\left|\frac{d}{dt}\langle v, \eta^t v'\rangle\right|_{A/\mathbb{R}} + \mathrm{tr}_{A/\mathbb{R}}\left(\overline{\langle v, \eta^t v'\rangle}\frac{d^2}{dt^2}\langle v, \eta^t v'\rangle\right).\end{aligned}$$
There are two distinct cases (see either [53, Section 3] or [43, Section 1]): η can be conjugated to a matrix of one of the following forms:
$$\begin{pmatrix} 1 & -\overline{a} & -|a|_{A/\mathbb{R}}/2 \\ 0 & 1_{n-1} & a \\ 0 & 0 & 1 \end{pmatrix}, a \in A^n \text{ or } \begin{pmatrix} 1 & 0 & b \\ 0 & 1_{n-1} & 0 \\ 0 & 0 & 1 \end{pmatrix}, b \in A \text{ totally imaginary.}$$
In the second case we get that $\frac{d^2}{dt^2}\eta^t = 0$, hence
$$\frac{d^2h}{dt^2} = 2\left|\frac{d}{dt}\langle v, \eta^t v'\rangle\right|_{A/\mathbb{R}} = 2|\overline{b}|_{A/\mathbb{R}} > 0.$$
In the first case (which we normalize so that $|a|_{A/\mathbb{R}} = 1$) we have at $t = 0$:
$$\frac{d^2h}{dt^2} = 2\left|\frac{d}{dt}\langle v, \eta^t v'\rangle\right|_{A/\mathbb{R}} - \mathrm{tr}_{A/\mathbb{R}}\left(\overline{v_{n+1}}v'_{n+1}\langle v, v'\rangle\right)$$
and hence the result follows from Lemma 4.1. □

4.3. Hyperbolic isometries and distance functions. In view of establishing the inequality (∗) in Lemma 2.1 axial isometries in negatively curved spaces have a much simpler behaviour than parabolic ones: one only needs to use the hyperbolicity of the space on which they act as soon as their minimal displacement is large enough, as was already observed in [10] (see Lemma 4.4 below). On the other hand, isometries with small enough minimal displacement which rotate non-trivially around their axis obviously do not satisfy (∗)

for all y; we study this phenomenon in more detail for real hyperbolic spaces below, obtaining an optimal criterion in Proposition 4.5.

4.3.1. *Gromov-hyperbolic spaces.* The following lemma is a slightly more precise version of Corollary 5.2 [10]. It has essentially the same proof; we will give the details, which are not contained in [10].

Lemma 4.4. *Let $\delta > 0$ and $d > 0$; there exists a constant $C(\delta, d)$ such that for any δ-hyperbolic space X and any axial isometry γ of X such that $\min(\gamma) \geq C(\delta, d)$ and any pair $(x, a) \in X$ we have*

$$\max(d(\gamma x, a), d(\gamma^{-1}x, a)) \geq d(x, a) + d.$$

Proof. Let γ be as in the statement (with the constant $C = C(\delta, d)$ to be determined later), let L be its axis. Let w, w', w'' be the projections of $x, \gamma x, \gamma^{-1}x$ on L, and v that of a. We will suppose (without loss of generality) that v lies on the ray in L originating at w and passing through w'.
Now let T be a metric tree with set of vertices constructed as follows: we take the geodesic segment on L containing all of w, w', w'' and v and we add the arcs $[x, w]$, etc. Then, for any two vertices u, u' of T we have

$$d_X(u, u') \leq d_T(u, u') \leq d_X(u, u') + c$$

where c depends only on δ (see the proof of Proposition 6.7 in [11]). In this tree we have

$$\begin{aligned} d_T(a, \gamma^{-1}x) &= d_X(a, v) + d_X(v, w) + d_X(w, w'') + d_X(w'', \gamma^{-1}x) \\ &= d_X(w, \gamma^{-1}w) + d_X(a, v) + d_X(v, w) + d_X(w, x) \\ &= \min(\gamma) + d_T(a, x) \end{aligned}$$

and using both inequalities above we get that

$$d_X(a, \gamma^{-1}x) \geq d_T(a, \gamma^{-1}x) - c \geq d_X(a, x) + \min(\gamma) - c.$$

We see that for $\min(\gamma) \geq C(\delta, d) = c + d$ the desired result follows. □

4.3.2. *A more precise result in real hyperbolic spaces.* We briefly discuss a quantitative version of Lemma 4.4. Bowditch observed (cf. Thm. 5.3 in [10]) that a group Γ which acts freely by axial transformations on the hyperbolic space $\mathbb{H}^n_{\mathbb{R}}$ is diffuse if every $\gamma \in \Gamma \setminus \{1\}$ has translation length at least $2\log(1 + \sqrt{2})$. We obtain a slight improvement relating the lower bound on the translation length more closely to the eigenvalues of the rotational part of the transformation.
Our proof is based on a calculation in the upper half-space model of $\mathbb{H}^n_{\mathbb{R}}$, i.e. we consider $\mathbb{H}^n_{\mathbb{R}} = \{ x \in \mathbb{R}^n \mid x_n > 0 \}$ with the hyperbolic metric d (see §4.6 in [56]). Every axial transformation γ on $\mathbb{H}^n_{\mathbb{R}}$ is conjugate to a transformation of the form $x \mapsto kAx$ where A is an orthogonal matrix in $\mathrm{O}(n-1)$ (acting on the first $n-1$ components) and $k > 1$ is a real number (see Thm. 4.7.4 in [56]). We say that A is the *rotational part* of γ. The translation length of γ is given by $\min(\gamma) = \log(k)$. We define the *absolute rotation* r_γ of γ to be the maximal

value of $|\lambda - 1|$ where λ runs through all eigenvalues of A. In other words, r_γ is merely the operator norm of the matrix $A - 1$. The absolute rotation measures how close the eigenvalues get to -1. It is apparent from Bowditch's proof that the case of eigenvalue -1 (rotation of angle π) is the problematic case whereas the situation should improve significantly for rotation bounded away from angle π. We prove the following sharp result.

Proposition 4.5. *An axial transformation γ of $\mathbb{H}^n_{\mathbb{R}}$ has the property*

$$(\bigstar) \qquad \max(d(x,\gamma y), d(x, \gamma^{-1}y)) > d(x,y) \text{ for all } x, y \in \mathbb{H}^n_{\mathbb{R}}$$

if and only if the translation length $\min(\gamma)$ *satisfies*

$$(\clubsuit) \qquad \min(\gamma) \geq \operatorname{arcosh}(1 + r_\gamma).$$

Using the same argument as above we immediately obtain the following improvement of Bowditch's Theorem 5.3 (we use Proposition 4.2 to take care of the unipotent elements).

Corollary. Let Γ be a group which acts freely by axial or unipotent transformations of the hyperbolic space $\mathbb{H}^n_{\mathbb{R}}$. If the translation length of every axial $\gamma \in \Gamma$ satisfies inequality ($\clubsuit$), then Γ is diffuse.

Remark. (1) It is a trivial matter to see that the converse of the corollary does not hold. Take any axial transformation $\gamma \neq 1$ which does not obey inequality ($\clubsuit$), then the diffuse group $\Gamma = \mathbb{Z}$ acts via γ on $\mathbb{H}^n$.
(2) If $\gamma \in \mathrm{SL}_2(\mathbb{C})$ is hyperbolic, with an eigenvalue $\lambda = e^{\ell/2}e^{i\theta/2}$ then the condition ($\clubsuit$) is equivalent to

$$\cosh(\ell) \geq 1 + \sqrt{2 - 2\cos(\theta)}.$$

Proof of Proposition 4.5. Let γ be an axial transformation which satisfies ($\clubsuit$). We will show that for all $x, y \in \mathbb{H}^n_{\mathbb{R}}$ we have $\max(d(x,\gamma y), d(x,\gamma^{-1}y)) > d(x,y)$. After conjugation we can assume that $\gamma(a) = Aka$ with $k > 1$ and $A \in \mathrm{O}(n-1)$. We take x, y to lie in the upper half-space model, then we may consider them as elements of $\mathbb{R}^n$. We will suppose in the sequel that $\|x\| \leq \|y\|$ in the euclidean metric of $\mathbb{R}^n$, and under this hypothesis we shall prove that $d(x,\gamma y) > d(x,y)$. If the opposite inequality $\|x\| \geq \|y\|$ holds we get that $d(y,\gamma x) > d(x,y)$, hence $d(x,\gamma^{-1}y) > d(x,y)$ which implies the proposition.
Using the definition of the hyperbolic metric and the monotonicity of cosh on positive numbers, it suffices to show

$$\|x - Aky\|^2 > k\|x - y\|^2.$$

In other words, we need to show that the largest real zero of the quadratic function

$$f(t) = t^2\|y\|^2 - t(\|x\|^2 + \|y\|^2 + 2\langle x, Ay - y\rangle) + \|x\|^2$$

is smaller than $\exp(\operatorname{arcosh}(1 + r_\gamma)) = 1 + r_\gamma + \sqrt{r_\gamma^2 + 2r_\gamma}$. We may divide by $\|y\|^2$ and we can thus assume $\|y\| = 1$ and $0 < \|x\| \leq 1$. The large root of $f(t)$

is

$$t_0 = \frac{\|x\|^2+1}{2} + \langle x, Ay-y\rangle + \frac{1}{2}\sqrt{(\|x\|^2+1+2\langle x, Ay-y\rangle)^2 - 4\|x\|^2}.$$

Note that if $r_\gamma = 0$, then $k > 1 = t_0$.
Suppose that $r_\gamma > 0$. Indeed, by Cauchy-Schwarz $|\langle x, Ay-y\rangle| < r_\gamma\|x\|$ and the inequality is strict since $x_n > 0$. As a consequence $t_0 < t(\|x\|)$ where

$$t(s) = \frac{s^2+1}{2} + r_\gamma s + \frac{1}{2}\sqrt{(s^2+1+2r_\gamma s)^2 - 4s^2}.$$

Finally, we determine the maximum of the function $t(s)$ for $s \in [0,1]$. A simple calculation shows that there is no local maximum in the interval $[0,1]$. We conclude that the maximal value is attained at $s = 1$ and is precisely

$$t(1) = 1 + r_\gamma + \sqrt{r_\gamma^2 + 2r_\gamma}.$$

Conversely, assume that (♣) does not hold. In this case we have $1 < k < 1 + r_\gamma + \sqrt{r_\gamma^2 + 2r_\gamma}$ and thus $r_\gamma \neq 0$. Choose some vector $y \in \mathbb{R}^n$ with $y_n = 0$ and $\|y\| = 1$ so that $\|Ay - y\| = r_\gamma$ (this is possible since r_γ is the operator norm of $A - 1$). We define $x = r_\gamma^{-1}(Ay - y)$ and we observe that $x \neq y$ since the orthogonal matrix A has no eigenvalues of absolute value exceeding one. The following inequalities hold:

$$\frac{\|x - k^{-1}A^{-1}y\|^2}{k^{-1}} \leq \frac{\|x - kAy\|^2}{k} < \|x - y\|^2.$$

The first follows from $\langle x, A^{-1}y\rangle \leq \langle x, y\rangle + r_\gamma = \langle x, Ay\rangle$. The second inequality follows from the assumption $k < 1 + r_\gamma + \sqrt{r_\gamma^2 + 2r_\gamma}$. Since the last inequality is strict, we can use continuity to find distinct x' and y' in the upper half-space (close to x and y), so that still

$$\max\left\{\frac{\|x' - k^{-1}A^{-1}y'\|^2}{k^{-1}}, \frac{\|x' - kAy'\|^2}{k}\right\} < \|x' - y'\|^2.$$

Interpreting x' and y' as points in the hyperbolic space, the assertion follows from the definition of the hyperbolic metric. □

4.4. Geometric finiteness. There are numerous equivalent definitions of geometric finiteness for discrete subgroups of isometries of rank one spaces, see for example [49, Section 3.1] or [56, Section 12.4] for real hyperbolic spaces. We shall use the equivalent definitions given by B. Bowditch in [9] for general negatively-curved manifolds.
The only facts from the theory of geometrically finite groups we will need in this section are the following two lemmas which are quite immediate consequences of the equivalent definitions.
In the rest of this section we will always use the following notation: whenever P is a parabolic subgroup in a rank-one Lie group and we write

$$P = MAN$$

this means that A is a split torus, M is compact and N is the unipotent radical of P (such a decomposition is essentially—up to conjugation of A and M by an element of N—unique).

Lemma 4.6. *Let G be a rank-one Lie group and $\Gamma \leq G$ be a geometrically finite subgroup, all of whose parabolic elements have finite-order eigenvalues. Then there is a subgroup $\Gamma' \leq \Gamma$ of finite index such that all parabolic isometries contained in Γ' are unipotent elements of G.*

Proof. From [9, Corollary 6.5] we know that Γ has only finitely many conjugacy classes of maximal parabolic subgroups; by residual finiteness of Γ we will be finished if we can show that for any parabolic subgroup P of G such that the fixed point of P in $\partial\mathbb{H}^n_{\mathbb{R}}$ is a cusp point, the group $\Lambda = \Gamma \cap P$ is virtually unipotent. Writing $P = MAN$ we see that it suffices to verify that the projection of Λ on A is trivial (Indeed, since then Λ is contained in MN, and its projection to M is finite because it has only finite-order elements by the hypothesis on eigenvalues, and it is finitely generated by [9, Proposition 4.1]). This follows from discreteness of Γ: if it contained an element λ with a non-trivial projection on A, then for any non-trivial $n \in N$ we have that either $\lambda^k n \lambda^{-k}$ or $\lambda^{-k} n \lambda^k$ goes to the identity of G; but since the fixed point of P is a cusp point for Γ the intersection $\Gamma \cap N$ must be nontrivial, hence there cannot exist such a λ. $\square$

Lemma 4.7. *Let G be a rank-one Lie group, Γ a torsion-free geometrically finite subgroup of G and $M_\Gamma = \Gamma\backslash X$. Then for any ℓ_0 there are only finitely many closed geodesics of length less than ℓ_0 in M_Γ.*

Proof. See also [56, Theorem 12.7.8]. One of Bowditch's characterizations of geometrical finiteness is the following: let $L_\Gamma \subset \partial X$ be the limit set of Γ, i.e. the closure of the set of points fixed by some nontrivial element of Γ, and let $Y_\Gamma \subset X$ be the convex hull in X of L_Γ. Let $C_\Gamma = \Gamma\backslash Y_\Gamma$ (the 'convex core' of M_Γ), and let $M_{[\varepsilon,+\infty[}$ be the ε-thick part of M_Γ. Then Γ is geometrically finite if and only if $C_\Gamma \cap M_{[\varepsilon,+\infty[}$ is compact (for some or any ε): see [9, Section 5.3]. It is a well-known consequence of Margulis' lemma that there is an $\varepsilon_0 > 0$ such that all geodesics in M_Γ of length less than ℓ_0 are contained in the ε_0-thick part. On the other hand it is clear that any closed geodesic of M_Γ is contained in C_Γ (since the endpoints of any lift are in L_Γ) and hence all closed geodesics of M_Γ with length $\leq \ell_0$ are contained in the compact set $C_\Gamma \cap M_{[\varepsilon_0,+\infty[}$, which implies that there are only finitely many such. $\square$

4.5. Main results.

4.5.1. *Action on the space.*

Theorem 4.8. *Let G be one of the Lie groups $\mathrm{SO}(n,1)$, $\mathrm{SU}(n,1)$ or $\mathrm{Sp}(n,1)$, X the associated symmetric space and let Γ be a geometrically finite subgroup of G. Suppose that all eigenvalues of parabolic elements of Γ are roots of unity. Then there exists a finite-index subgroup $\Gamma' \subset \Gamma$ such that Γ' acts diffusely on X.*

Proof. Let Γ' be a finite-index subgroup of Γ such that all semisimple elements $\gamma \in \Gamma'$ have $\min(\gamma) > C(\delta_X, 1)$ (where δ_X is a hyperbolicity constant for X, which is Gromov-hyperbolic since it is a negatively-curved, simply connected Riemannian manifold, and $C(\delta_X, 1)$ is the constant from Lemma 4.4)—such a subgroup exists by Lemma 4.7 and the residual finiteness of Γ. By Lemma 4.6 we may also suppose that the parabolic isometries in Γ' are exclusively unipotent.
Now we can check that the hypothesis $(*)$ in Lemma 2.1 holds for the action of Γ on X: for axial isometries we only have to apply Lemma 4.4, and for unipotent elements Proposition 4.2. □

The hypothesis on eigenvalues of parabolic elements is equivalent to asking that every parabolic subgroup of Γ contains a finite-index subgroup which consists of unipotent elements. It is necessary for an application of Lemma 2.1, as shown by the following construction.

Lemma 4.9. *For $n \geq 4$ there exists a discrete, two-generated free subgroup Γ of $\mathrm{SO}(n,1)$ such that for all $x \in \mathbb{H}^n_{\mathbb{R}}$ there is a $y \in \mathbb{H}^n_{\mathbb{R}}$ and a $g \in \Gamma \setminus \{1\}$ such that*

$$d(x,y) \geq d(gx,y),\ d(g^{-1}x,y).$$

Proof. It suffices to prove this lemma for $\mathrm{SO}(4,1)$. Let ω be an infinite-order rotation of $\mathbb{R}^2$ and let ϕ be the isometry of $\mathbb{R}^3 = \mathbb{R} \times \mathbb{R}^2$ given by $(t,x) \mapsto (t+1, \omega \cdot x)$. Then it is easy to see that for any k and any x not on the axis $\mathbb{R} \times 0$ of ϕ the bisectors between x and $\phi^{\pm k}x$ intersect. Let $\widetilde{\phi}$ be the isometry of $\mathbb{H}^4_{\mathbb{R}}$ obtained by taking the Poincaré extension of ϕ (i.e. we fix a point on $\partial\mathbb{H}^4_{\mathbb{R}}$ and define $\widetilde{\phi}$ by identifying the horospheres at this point with the Euclidean three–space on which ϕ acts), which will also not satisfy $(*)$ for all points outside of a two dimensional totally geodesic submanifold Y_ϕ.
Now take ϕ_1, ϕ_2 as above. There exists a $g \in \mathrm{Isom}(\mathbb{H}^4_{\mathbb{R}})$ such that $gY_{\phi_2}g^{-1} \cap Y_{\phi_2} = \emptyset$, and then for any $k_1, k_2 > 0$ the group $\langle \widetilde{\phi}_1^{k_1}, \widetilde{\phi}_2^{k_2} \rangle$ satisfies the second conclusion of the lemma. It remains to prove that for k_1, k_2 large enough it is a discrete (and free) group. This is done by a very standard argument which goes as follows: There are disjoint open neighbourhoods U_i of $\mathrm{Fix}(\widetilde{\phi}_i)$ in $\partial\mathbb{H}^4_{\mathbb{R}}$ (not containing $\mathrm{Fix}(\widetilde{\phi}_j), j \neq i$) and positive integers k_1, k_2 such that for all $k \in \mathbb{Z}, |k| \geq k_i$ we have $\widetilde{\phi}_i^k(\mathbb{H}^4_{\mathbb{R}} \setminus U_i) \subset U_i$. Now we can apply the ping-pong lemma of Klein to obtain freeness and discreteness of $\langle \widetilde{\phi}_1^{k_1}, \widetilde{\phi}_2^{k_2} \rangle$: fix a $\xi \in \partial\mathbb{H}^4_{\mathbb{R}} \setminus (U_1 \cup U_2)$, then any non-trivial reduced word in $\widetilde{\phi}_1, \widetilde{\phi}_2$ sends ξ inside one of U_1 or U_2, hence the orbit of ξ is discrete in $\partial\mathbb{H}^4_{\mathbb{R}}$ (proving discreteness of $\langle \widetilde{\phi}_1^{k_1}, \widetilde{\phi}_2^{k_2} \rangle$) and any such word is nontrivial in $\mathrm{SO}(4,1)$ (proving freeness). □

On the other hand this phenomenon cannot happen in $\mathbb{H}^2_{\mathbb{R}}, \mathbb{H}^3_{\mathbb{R}}$, which yields the following corollary of Theorem 4.8.

Corollary 1. If Γ is a finitely generated discrete subgroup of $\mathrm{SL}_2(\mathbb{C})$ then Γ is virtually diffuse.

Proof. Since in dimension three all Kleinian groups are isomorphic to geometrically finite ones (this is a consequence of Thurston's hyperbolization theorem for Haken manifolds, as explained in [49, Theorem 4.10]) the result would follow if we can prove diffuseness for the latter class. But parabolic isometries of $\mathbb{H}^3$ are necessarily unipotent (since if an element of $\mathrm{SL}_2(\mathbb{C})$ has two equal eigenvalues, they must be equal to ± 1, and hence it is unipotent in the adjoint representation), and thus we can apply Theorem 4.8 to deduce that a geometrically finite Kleinian group in dimension three has a finite-index subgroup which acts diffusely on $\mathbb{H}^3$.
We could also deduce Corollary 1 from the veracity of the Tameness conjecture [1], [17] and the virtual diffuseness of three–manifolds groups, Theorem D from the introduction. □

Also, when parabolic subgroups of Γ are large enough[3] the hypothesis should be satisfied. We will be content with the following application of this principle.

Corollary 2. If Γ is a lattice in one of the Lie groups $\mathrm{SO}(n,1)$, $\mathrm{SU}(n,1)$ or $\mathrm{Sp}(n,1)$ then Γ is virtually diffuse.

Proof. A lattice Γ in a rank one Lie group G is a geometrically finite group (cf. 5.4.2 in [9]), hence we need to prove that the parabolic isometries contained in Γ have only roots of unity as eigenvalues. In the case that Γ is arithmetic there is a quick argument: for any $\gamma \in \Gamma$, the eigenvalues of γ are algebraic numbers. If in addition γ is parabolic, then all its eigenvalues are of absolute value one as well as their conjugates (because the group defining Γ is compact at other infinite places). A theorem of Kronecker [27, Theorem 1.31] shows that any algebraic integer in $\mathbb{C}$ whose Galois conjugates are all of absolute value one must be a root of unity, and it follows that the eigenvalues of γ are roots of unity.
One can also use a more direct geometric argument to prove this in full generality. Let $P = MAN$ be a parabolic subgroup of G which contains a parabolic element of Γ; then it is well-known that $\Gamma \cap P$ is contained in MN (see the proof of Lemma 4.6 above). Also $\Lambda = \Gamma \cap N$ is a lattice in N, in particular $\Lambda \backslash N$ is compact (this follows from the Margulis Lemma [9, Proposition 3.5.1], which implies that horosphere quotients inject into $\Gamma \backslash X$, and the finiteness of the volume of $\Gamma \backslash X$). Corollary 2 will then follow from the next lemma.

Lemma 4.10. *Let N be a simply connected nilpotent Lie group containing a lattice Λ, and $Q \leq \mathrm{Aut}(N)$ a subgroup which preserves Λ, all of whose elements have only eigenvalues of absolute value one (in the representation on the Lie algebra $\mathfrak{n}$). Then these eigenvalues are in fact roots of unity.*

Proof. The exponential map $\exp : \mathfrak{n} \to N$ is a diffeomorphism. By [55, Theorem 2.12], there is a lattice L in the vector space $\mathfrak{n}$ such that $\langle \exp(L) \rangle = \Lambda$. It follows that the adjoint action of Q preserves L, hence for any $q \in Q$ the

[3]For example, in the real hyperbolic case, when their span in the Lie algebra is of codimension smaller than one.

characteristic polynomial of $\mathrm{Ad}(q)$ has integer coefficients, hence its eigenvalues are the conjugates of some finite set of algebraic integers. Since they are also all of absolute value one it follows from Kronecker's theorem that they must be roots of unity. □

It follows that, in the above setting, the image of $\Gamma \cap P$ in M has a finite-order image in $\mathrm{Aut}(N)$ where M acts by conjugation. This action is faithful (because an element of M cannot act trivially of an horosphere associated to N, otherwise it would act trivially on the whole of X since it preserves these horospheres) and it follows that the hypothesis on eigenvalues in Theorem 4.8 is satisfied by Γ. □

4.5.2. *Action on the boundary.*

Theorem 4.11. *Let Γ, G be as in the statement of Theorem 4.8. Then there is a finite-index $\Gamma' \subset \Gamma$ such that for any parabolic fixed point $\xi \in \partial X$ for Γ' with stabilizer Λ_ξ in Γ' the action of Γ' on Γ'/Λ_ξ is diffuse.*

Proof. We take a finite-index subgroup $\Gamma' \leq \Gamma$ as in the proof of Theorem 4.8 above. The key point is the following lemma.

Lemma 4.12. *There is a dense subset $S_{\Gamma'} \subset X$ such that for any $x_0 \in S_{\Gamma'}$ and any parabolic fixed point ξ of Γ', if b_ξ is a Busemann function at ξ we have*

$$\forall g \in \Gamma',\, g \notin \Lambda_\xi : \max\left(b_\xi(gx_0), b_\xi(g^{-1}x_0)\right) > b_\xi(x_0). \tag{3}$$

Proof. Fix ξ and b_ξ as in the statement. By definition of a Busemann function there is a unit speed geodesic ray $\sigma : [0, \infty[\to X$ running to ξ in $X \cup \partial X$, such that for all $x \in X$ we have

$$b_\xi(x) = \lim_{t \to +\infty} \left(d(x, \sigma(t)) - t\right).$$

On the other hand, by construction of Γ' (using Lemma 4.4) we know that for all axial isometries $g \in \Gamma' \setminus \{1\}$ we have

$$\forall t \geq 0 \ \max\left(d(gx_0, \sigma(t)), d(g^{-1}x_0, \sigma(t))\right) \geq d(x_0, \sigma(t)) + 1;$$

passing to the limit we obtain (3) for all such g and for any choice of x_0.
Now we show that for certain generic x_0 the same is true for unipotent isometries. In any case, for any unipotent isometry g of X, it follows from Proposition 4.2 and the same argument as above that

$$\max(b_\xi(g^{-1}x_0), b_\xi(gx_0)) \geq b_\xi(x_0) \tag{4}$$

for all x_0. We want to choose x_0 in order to be able to rule out equality if $g \in \Gamma' - \Lambda_\xi$. For a given unipotent isometry η and a $\zeta \in \partial X$ with $\eta\zeta \neq \zeta$ define

$$E_{\zeta,\eta} = \{x \in X \mid b_\zeta(\eta x) = b_\zeta(x)\}$$

(note that this does not depend on the choice of the Busemann function b_ζ). This is an embedded hyperplane in X, and hence (by Baire's theorem) the subset

$$S_{\Gamma'} = X - \bigcup_{\zeta,\eta} E_{\zeta,\eta}$$

where the union runs over all parabolic elements η of Γ' and all parabolic fixed points ζ of Γ', is dense in X. Moreover, by the definition of $S_{\Gamma'}$, for $x_0 \in S_{\Gamma'}$ we never have $b_\xi(gx_0) = b_\xi(x_0)$ for any unipotent $g \in \Gamma'$ with $g\xi \neq \xi$. Thus (4) has to be a strict inequality. □

Let $\xi_0 \in \partial X$ be a parabolic fixed point of Γ' and b_{ξ_0} a Busemann function at ξ_0. We write $\Lambda = \Lambda_{\xi_0}$. The function b_{ξ_0} is Λ-invariant and if we choose some $x_0 \in X$ we may define a function $f = f_{x_0}$ on Γ'/Λ by

$$f(\gamma\Lambda) = b_{\xi_0}(\gamma^{-1}x_0) = b_{\gamma\xi_0}(x_0). \tag{5}$$

By the lemma this function satisfies

$$\forall \gamma\Lambda \in \Gamma'/\Lambda,\ \forall g \in \Gamma',\ g \notin \gamma\Lambda\gamma^{-1} : \max\left(f(g\gamma\Lambda), f(g^{-1}\gamma\Lambda)\right) > f(\gamma\Lambda), \tag{6}$$

whenever $x_0 \in S_{\Gamma'}$. Indeed, we have

$$\max\left(f(g\gamma\Lambda), f(g^{-1}\gamma\Lambda)\right) = \max\left(b_{\gamma\xi_0}(gx_0), b_{\gamma\xi_0}(g^{-1}x_0)\right)$$

and according to (3) the right-hand side is strictly larger than $b_{\gamma\xi_0}(x_0) = f(\gamma\Lambda)$. The existence of a function f satisfying (6) implies that the action Γ' on Γ'/Λ is weakly diffuse, i.e. every non-empty finite subset $A \subset \Gamma'/\Lambda$ has at least *one* extremal point. Indeed, any $a \in A$ such that $f(a)$ realizes the maximum of f on A is extremal in A.
Using an additional trick we can actually deduce diffuseness. Let $A \subset \Gamma'/\Lambda$ be finite with $|A| \geq 2$, and let a be an extremal point. By shifting A we can assume that $a = \Lambda$. Now let ξ_0 be the fixed point of Λ and b_{ξ_0} a Busemann function. Choose $x_0 \in S_{\Gamma'}$ such that x_0 is (up to Λ) the only point realizing the minimum of b_{ξ_0} on $\Gamma' x_0$ (this is possible by taking x_0 in a sufficiently small horoball at ξ_0, since $S_{\Gamma'}$ is dense) and define f on Γ'/Λ as in (5). By construction f takes it's minimal value at a. So let $b \in A$ be a point where f takes a maximal value. By the given argument b is extremal in A. On the other hand $f(b) > f(a)$ and so $b \neq a$. We conclude that A has at least two extremal points. □

5. Fundamental groups of three–manifolds

In this section we prove Theorem D, whose statement we recall now :

Theorem. *Let M be a compact three–manifold and $\Gamma = \pi_1(M)$ its fundamental group. Then there is a finite-index subgroup $\Gamma' \leq \Gamma$ which is diffuse.*

The proof is a rather typical application of Geometrization. We begin with an algebraic result on graph products, afterwards we use it to construct a suitable covering (cf. [37]).

5.1. Algebraic preliminaries: a gluing lemma. Bowditch [10] showed that if Γ is the fundamental group of a graph of groups such that for any vertex group Γ_i and adjacent edge group Λ_i, both the group Λ_i and the action of Γ_i on Γ_i/Λ_i are diffuse, then Γ is diffuse. In order to glue manifolds it is necessary to understand graph products of virtually diffuse groups. For free products there is a very simple argument.

Lemma 5.1. *The free product $G = G_1 * G_2$ of two virtually diffuse groups G_1 and G_2 is again virtually diffuse.*

Proof. Let $H_i \leq_f G_i$ be a finite index diffuse subgroup. Consider the homomorphism $\phi : G \to G_1 \times G_2$. The kernel K of ϕ is a free group (cf. I. Prop. 4 in [60]). Let H denote the inverse image of $H_1 \times H_2$ under ϕ. The subgroup H has finite index in G and $H \cap K$ is a free group. We get a short exact sequence

$$1 \longrightarrow K \cap H \longrightarrow H \to H_1 \times H_2 \longrightarrow 1.$$

From Theorem 1.2 in [10] we see that H is diffuse. □

Note that the same argument shows that the free product of diffuse groups is diffuse. In order to understand amalgamated products and HNN extensions of virtually diffuse groups one needs to argue more carefully.
We will use the Bass-Serre theory of graph products of groups. We shall use the notation of [60]. Recall that a graph of groups (G, Y) is a *finite* graph Y with vertices $V(Y)$ and edges $E(Y)$. Every edge e has an origin $o(e) \in V(Y)$ and a terminus $t(e) \in V(Y)$. Moreover for every edge there is an opposite edge $\bar{e}$. To every vertex $P \in V(Y)$ and every edge $e \in E(Y)$ there are attached groups G_P and $G_e = G_{\bar{e}}$. Moreover, for every edge e there is a monomorphism $i_e : G_e \to G_{t(e)}$ usually denoted by $a \mapsto a^e$. To a graph of groups one attaches a fundamental group $\pi_1(G, Y)$ – the graph product.
Let (G, Y) be a graph of groups. A *normal subcollection* (N, Y) consists of two families $(N_P \trianglelefteq G_P)_{P \in V(Y)}$ and $(N_e \trianglelefteq G_e)_{e \in E(Y)}$ of normal subgroups in the vertex and edge groups which are *compatible* in the sense that

$$i_e(N_e) = i_e(G_e) \cap N_{t(e)} \quad \text{and} \quad N_e = N_{\bar{e}}$$

for every edge $e \in E(Y)$. We say that (N, Y) is of finite index, if for every vertex P the index of N_P in G_P is finite.

Lemma 5.2. *Let (G, Y) be a graph of finite groups. The fundamental group $\Gamma = \pi_1(G, Y)$ is residually finite and virtually free.*

Proof. The residual finiteness follows from Theorem 3.1 of Hempel [37]. To apply his result we need to specify sufficiently small normal subcollections (H, Y) in (G, Y) such for every $P \in V(Y)$ the group H_P has finite index in G_P. Since we are dealing with finite groups it is easy to check that we can simply choose $H_P = \{1\}$ and $H_e = \{1\}$ for every vertex P and edge e.
Using that Γ is residually finite, we can find a finite index subgroup $N \trianglelefteq \Gamma$ which intersects the embedded vertex group G_P trivially for any of the finitely many vertices $P \in V(Y)$. Therefore, the subgroup N acts freely (without edge inversion) on the Bass-Serre tree associated with the graph (G, Y). We deduce that N is a free group [60, I. Thm. 4]. □

Let (G, Y) be a graph of groups and let (N, Y) be a normal subcollection. To such a data we can associate a quotient graph of groups (H, Y) where $H_P = G_P/N_P$ and $H_e = G_e/N_e$ for all vertices P and edges e. There is a natural surjective quotient morphism $q : \pi_1(G, Y) \to \pi_1(H, Y)$. We are now able to state and prove the (algebraic) gluing Lemma.

Lemma 5.3 (Gluing Lemma). *Let (G, Y) be a graph of groups such that*
- *(i) every edge group G_e is diffuse*
- *(ii) there is a normal subcollection (N, Y) of finite index such that for every edge $e \in E(Y)$ the group $N_{t(e)}$ acts diffusely on $G_{t(e)}/i_e(G_e)$.*

In this case the fundamental group $\Gamma = \pi_1(G, Y)$ is virtually diffuse.

Proof. Consider the associated quotient morphism $q : \Gamma \to \pi_1(H, Y)$. The kernel $\mathfrak{N}$ of q is the normal subgroup generated by the groups $(N_P)_{P\in V(Y)}$. Let Γ and $\mathfrak{N}$ act on the Bass-Serre tree T associated with (G, Y). The stabilizer in Γ (resp. $\mathfrak{N}$) of a vertex $v \in V(T)$ above $P \in V(Y)$ is isomorphic to G_P (resp. N_P). It acts on the set of adjacent edges $E(v) \subset E(T)$. Pick an edge $e \in E(Y)$ with $t(e) = P$. As a set with G_P action $E(v)$ is isomorphic to $G_P/i_e(G_e)$. By assumption (ii) the action of N_P on $G_P/i_e(G_e)$ is diffuse. By a result of Bowditch [10, Prop. 2.2] we deduce that $E(T)$ is a diffuse $\mathfrak{N}$ set. Since the edge groups are assumed to be diffuse, we see that $\mathfrak{N}$ is diffuse.
The quotient (H, Y) is a graph of finite groups, we know from Lemma 5.2 that it is virtually free. Since free groups are diffuse, the short exact sequence

$$1 \longrightarrow \mathfrak{N} \longrightarrow \Gamma \longrightarrow \pi_1(H, Y) \longrightarrow 1$$

implies the assertion by Thm. 1.2 (2) of [10]. □

5.2. Geometrization and the proof of Theorem D.

5.2.1. *Definitions.* We recall here the definitions which allow to state the Geometrization Theorem which was conjectured by W. Thurston ([64], see also [59]) and proven by G. Perelman [51, 52] (see also [44] for a complete account of Perelman's proof).
In the following we consider (without loss of generality) only orientable manifolds. A three–manifold M is called *irreducible* if all embedded 2-spheres in M bound a ball. A manifold is *prime* if it is irreducible or homeomorphic to $S^1 \times S^2$. According to the Kneser–Milnor decomposition every closed three–manifold is a finite connected sum of prime manifolds. A closed irreducible manifold M has a further topological decomposition, called the Jaco–Shalen–Johansson decomposition, which consists in a canonical collection of embedded, essentially disjoint 2-tori in M (see [41]). The Geometrization Theorem states that every connected component of the complement in M of this collection of tori is either a finite volume hyperbolic manifold or Seifert fibered.

5.2.2. *Virtual diffuseness.* The following lemma treats the pieces of the Geometrization Theorem. It is the key ingredient for Theorem D.

Lemma 5.4. *Let M be a compact three–manifold with incompressible toric boundary. If M is either hyperbolic of finite volume or Seifert fibered, then $\Gamma = \pi_1(M)$ contains a diffuse subgroup Γ' of finite index. Moreover, if M has* *non-empty boundary, then for almost all primes p the group Γ' can be chosen so that for any peripheral subgroup Λ of Γ*

(a) the Γ'-action on Γ/Λ is diffuse and
(b) $\Gamma' \cap \Lambda$ is the characteristic subgroup of index p^2.

Proof. Assume first that M is closed. If M is Seifert fibered, then $\pi_1(M)$ is a an extension of a group which is virtually a surface group by a cyclic group C (cf. Lemma 3.2 in [59]). If C is infinite, such a group is virtually diffuse by the results of Bowditch [10]. Otherwise M is covered by S^3 and the fundamental group is finite. If M is hyperbolic, then the virtual diffuseness follows from Theorem B.
Now we turn to the case where M has non-empty boundary. Assume first that M is hyperbolic. In $\pi_1(M)$ there are only finitely many, say m, hyperbolic conjugacy classes represented by elements $h_1, \dots, h_m$ with translation length less than $2\log(1+\sqrt{2})$ (cf. Lemma 4.7). By Lemma 4.1 of [37] we can find, for almost all primes p, a normal subgroup of finite index $\Gamma'_p \leq \pi_1(M)$ which does not contain $h_1, \dots, h_m$ and which intersects each peripheral subgroup in its characteristic subgroup of index p^2. Using Theorem 4.11 such a group Γ'_p is diffuse and acts diffusely on Γ/Λ for any peripheral Λ when p is large enough.
Finally assume that M is Seifert fibered. There is a short exact sequence

$$1 \longrightarrow \mathbb{Z} \longrightarrow \pi_1(M) \xrightarrow{q} G \longrightarrow 1$$

where $\mathbb{Z}$ is generated by the regular fibers and G is the fundamental group of a two dimensional orbifold B with non-empty boundary. Taking the finite index subgroup of elements commuting with the regular fibres (which contains the peripheral subgroups), we can assume that the extension is central. Since the boundary of M is incompressible, the simple closed boundary curves $d_1, \dots d_k$ of B have infinite order in G. For almost all primes p there is a free normal subgroup $G_p \subset G$ of finite index such that $G_p \cap \langle d_i \rangle = \langle d_i^p \rangle$. One way to see this is to argue using the presentation of G as given in [36, 12.1]. Geometrically this can be seen as follows: Glue a disc with a p-cone point into every boundary curve of B. For almost all p the resulting orbifold B_p is good and has hence a finite sheeted regular cover $\widetilde{B_p}$ which is a manifold. Removing the inverse images of the glued discs we obtain a finite covering space S_p of B which is a compact surface so that the boundary components are p-fold covers of the boundary components of B. Since a compact surface with non-empty boundary has a free fundamental group the claim follows.
The finite sheeted cover $\widetilde{M_p}$ corresponding to $q^{-1}(G_p)$ has fundamental group isomorphic to $\mathbb{Z} \times G_p$. Finally the group $\Gamma'_p = p\mathbb{Z} \times G_p$ is diffuse and intersects the peripheral subgroups in their characteristic subgroups of index p^2. It remains to verify that the action of Γ'_p on Γ/Λ is diffuse. This action factors through the group G_p and so the assertion follows, for p large enough, from Theorem 4.11 if we embed G_p as a discrete subgroup into $\mathrm{SL}_2(\mathbb{R})$.

There is another argument for the diffuseness of this action: We can assume that the surface S_p has more than one boundary component, and so the boundary curves can be chosen to be part of a free generating set. Let F be a free group and $f \in F$ an element of a free generating set, then by Prop. 2.2 in [10] the action of F on $F/\langle f \rangle$ is diffuse. □

5.2.3. *Proof of Theorem D.* Let M be a compact three–manifold; by doubling it (and since virtual diffuseness passes to subgroups) we may assume that it is in fact closed. By Lemma 5.1 and the Kneser–Milnor decomposition we may assume that M is irreducible. An irreducible manifold admits a geometric decomposition (see 5.2.1), which yields a decomposition of $\pi_1(M)$ as a graph of groups whose vertex groups are fundamental groups of Seifert fibered or hyperbolic manifolds and the edge groups are peripheral subgroups. Choosing a prime number p which is admissible for all the occurring pieces, it follows from Lemma 5.4 that this graph of groups has a normal subcollection which satisfies the hypotheses of Lemma 5.3.

5.3. Three–dimensional infra-solvmanifolds. A three–dimensional *solvmanifold* is a (left) quotient of the solvable Lie group

$$\mathbf{Sol} = \mathbb{R}^2 \rtimes \mathbb{R}; \quad t \cdot x = \begin{pmatrix} e^t & \\ & e^{-t} \end{pmatrix} \cdot x$$

by a discrete subgroup; an *infra-solvmanifold* is a quotient of such by a finite group acting freely. Any left-invariant Riemannian metric on **Sol** induces a complete Riemannian metric on an infra-solvmanifold. A compact solvmanifold is finitely covered by a torus bundle (see for example [59, Theorem 5.3 (i)]), hence its fundamental group contains a subgroup of finite index which is an extension of $\mathbb{Z}^2$ by $\mathbb{Z}$. More precisely, this group will be isomorphic to some

$$\Gamma_A = \langle \mathbb{Z}^2, t \mid \forall v \in \mathbb{Z}^2, tvt^{-1} = Av \rangle$$

where $A \in \mathrm{SL}_2(\mathbb{Z})$ is not unipotent. Such a group is diffuse by [10, Thm 1.2]. On the other hand we will now explain how to construct infra-solvmanifolds (so-called 'torus semi-bundles') of dimension three with zero first Betti number (by gluing I-bundles over Klein bottles, see [35]), which are then not locally indicable and hence not diffuse.
The following result is a special case of Proposition 3.11. We shall give another geometric argument (see also [33, Corollary 8.3] for a complete description of the groups of isometries acting properly discontinuously, freely and cocompactly on **Sol** from which it follows easily).

Proposition 5.5. *In every commensurability class of compact three–dimensional infra-solvmanifolds there is a manifold with non-diffuse fundamental group.*

Proof. In this proof we will first describe a topological construction from [35] of sol-manifolds with $b_1 = 0$, and then show that any sol-manifold is commensurable to one of these.

Let N be the non-trivial I-bundle over the Klein bottle, so that $\partial N = \mathbb{T}^2$. Then for any mapping class $B \in \mathrm{SL}_2(\mathbb{Z})$ of $\mathbb{T}^2$ the gluing $M = N \cup_\phi N$ has $b_1(M) = 0$ or is Seifert; in the former case it is a sol-manifold and is doubly covered by the torus bundle with holonomy $A_0 = SB^{-1}SB$ where S is the symmetry $(x, y) \mapsto (-x, y)$. In this way we get all A_0s of the form

$$A_0 = \begin{pmatrix} b & 2a \\ 2c & b \end{pmatrix} \tag{7}$$

where $a, b, c \in \mathbb{Z}$: this follows from a direct computation.
On the other hand we will see that for any hyperbolic $A \in \mathrm{SL}_2(\mathbb{Z})$ there is an integer $n > 0$ such that A^n is conjugated to a matrix of the form above. This implies the proposition since then the mapping torus of A^n (which covers that of A) has a quotient with $b_1 = 0$. Let us prove this claim: take L to be the geodesic line (in the Poincaré upper half-plane) orthogonal to the axis of A ending at ∞; then we can find $h \in \mathrm{SL}_2(\mathbb{Q})$ such that $hL = (0, \infty)$. Since h commensurates $\mathrm{SL}_2(\mathbb{Z})$ the group $hA^{\mathbb{Z}}h^{-1} \cap \mathrm{SL}_2(\mathbb{Z})$ is non-trivial ; take any $A' \neq \mathrm{Id}$ in there, then A' has both diagonal coefficients equal (this also follows from a simple computation). Taking A_0 to be the cube or square of A' we get a matrix of the form (7) above. $\square$

Appendix A. Computational aspects

A.1. Finding ravels. Given a group Γ it is a substantial problem to decide whether or not the group is diffuse. To a certain degree this problem is vulnerable to a computational approach which will be explained in this section.
For all the following algorithms we suppose that we have a way of solving the word problem in a given group Γ; in practice we used computations with matrices to do this. We will not make reference to the group Γ in the algorithms.
The first algorithm determines, given a finite subset A of Γ and an element $a \in A$, whether a is extremal in A or not.

Algorithm 1 Given $a \in A \subset \Gamma$, determines if a is extremal in A

function IsExtremal(a, A)
  $B = A \setminus \{a\}$
  for all $b \in B$ do
    if $ab^{-1}a \in A$ then return False ▷ If $b = ga$ and $g^{-1}a = ab^{-1}a \in A$ then a is not extremal.
  return True

The following algorithm returns the largest ravel contained in A by successively removing extremal points. If A contains no ravel, then it returns the empty set. Of course, the algorithm is not able to decide if a ravel exists at all (hence is of no use to prove that a group is diffuse).

Algorithm 2 Given $A \subset \Gamma$, finds the largest ravel contained in A

```
function FindRavel(A)
    for all a ∈ A do
        if IsExtremal(a, A) = True then return FindRavel(A \ {a})
    return A        ▷ No extremal point was found in A, so A is a ravel or empty
```

Finally, it may be of interest to determine minimal ravels; the following algorithm, starting from a ravel A, finds a minimal one contained in A (note that the result may depend on the order on which the elements of A are looped over).

Algorithm 3 Given a ravel $A \subset \Gamma$, finds a minimal ravel contained in A

```
function MinRavel(A)
    for all a ∈ A do
        B = FindRavel(A \ {a})
        if B ≠ ∅ then return MinRavel(B)
    return A
```

To prove Proposition 2.3 we ran (with two different implementations in Magma [8] and in Sage/Python [62]) the algorithms to test diffuseness on the group with presentation

$$\langle a, b | a^2b^2a^2b^{-1}ab^{-1},\ a^2b^2a^{-1}ba^{-1}b^2 \rangle,$$

which is the fundamental group of the Weeks manifold, the hyperbolic three–manifold of smallest volume. We actually used the representation to $\mathrm{SL}_2(\mathbb{C})$ given in the proof of Proposition 3.2 in [20]:

$$a = \begin{pmatrix} x & 1 \\ 0 & x^{-1} \end{pmatrix}, \quad b = \begin{pmatrix} x & 0 \\ 2-(x+x^{-1}) & x^{-1} \end{pmatrix}$$

where

$$x^6 + 2x^4 - x^3 + 2x^2 + 1 = 0.$$

It turns out that the word metric ball of radius four in the generators a, b contains a ravel of cardinality 141 (further computation showed that the latter contains a minimal ravel of cardinality 23).

A.2. Implementation.

A.2.1. *SAGE.* The Sage implementation of the algorithm (for linear groups) can be found in [26]. It has to be run in a Sage environment, and the main function is `max_diff`, which takes as input a pair (S, M) where M is a Sage MatrixSpace object, and S a collection of invertible matrices in M. Its output is the (possibly empty) maximal ravel contained in S. The file also contains the function `ball`, which inputs a triple $(r, \mathrm{gens}, \mathrm{M})$ which computes the ball of radius r in the group generated by the set gens of invertible matrices in M (in the word metric associated to gens). Another file in [26] can be run directly in a Sage environment and outputs a ravel of cardinality 141 in the Weeks manifold group.

A.2.2. *MAGMA.* An implementation for the MAGMA computer algebra system can be found in [26]. It includes functions `findRavel`, `findMinRavel` and a procedure `BallWeeks` to generate a ball of given radius in the Weeks manifold group. To compute a ravel in the Weeks manifold group run the following lines

```
|> B := BallWeeks(4);
|> findRavel(B);
```

Appendix B. A diffuse group which is not left-orderable by Nathan M. Dunfield

This appendix is devoted to the proof of

Theorem B.1. *Let N be the closed orientable hyperbolic 3-manifold defined below. Then $\pi_1(N)$ is diffuse but not left-orderable.*

This example was found by searching through the towers of finite covers of hyperbolic 3-manifolds studied in [18, §6]. There, each manifold has $b_1 = 0$ (which is necessary for π_1 to be non–left-orderable) and the length of the systole goes to infinity (so that we can apply Bowditch's criterion for diffuseness). We begin by giving two descriptions of N, one purely arithmetic and the other purely topological.

B.1. Arithmetic description. Throughout this section, a good reference for arithmetic hyperbolic 3-manifolds is [48]. Let $K = \mathbb{Q}(\alpha)$ be the number field where $\alpha^3+\alpha-1 = 0$; this is the unique cubic field with discriminant -31. It has one real embedding and one pair of complex embeddings; our convention is that the complex place corresponds to $\alpha \approx -0.3411639 + 1.1615414i$. Its integer ring $\mathcal{O}_K$ has unique factorization, so we will not distinguish between prime elements and prime ideals of $\mathcal{O}_K$. The unique prime of norm 3 in $\mathcal{O}_K$ is $\pi = \alpha + 1$, and let D be the quaternion algebra over K ramified at exactly π and the real place of K. Concretely, we can take D to be generated by i and j where $i^2 = -1$, $j^2 = -3$ and $k = ij = -ji$. The manifold N will be the congruence arithmetic hyperbolic 3-manifold associated to D and the level π^3, whose detailed construction we now give.
Let $\mathcal{O}_D$ be a maximal order in D; this is unique up to conjugation by [48, Example 6.7.9(3)]. Let $\mathcal{O}_D^1$ denote the elements of $\mathcal{O}_D$ of (reduced) norm 1. At the complex place of K, the algebra $\mathbb{C} \otimes_K D$ is just the matrix algebra $M_2(\mathbb{C})$. Let Λ be the subgroup of $\mathrm{PSL}_2(\mathbb{C}) \cong \mathrm{Isom}^+ \mathbb{H}^3$ which is the image of $\mathcal{O}_D^1$ under the induced map $D^1 \to \mathrm{SL}_2(\mathbb{C}) \to \mathrm{PSL}_2(\mathbb{C})$. Since D is a division algebra, Λ is a cocompact lattice. Let K_π be the π-adic completion of K, which is isomorphic to $\mathbb{Q}_3$. Let $D_\pi = K_\pi \otimes_K D$, which is the unique quaternion division algebra over K_π [48, §2.6]. Define $w\colon D_\pi \to \mathbb{Z}$ by $w = \nu \circ n$ where

$\nu \colon K_\pi \to \mathbb{Z}$ is the (logarithmic) valuation and $n \colon D_\pi \to K_\pi$ is the norm function. Then $\mathcal{O}_\pi = \{u \in D_\pi \mid w(u) \geq 0\}$ is the valuation ring of D_π and $\mathcal{Q} = \{u \in D_\pi \mid w(u) \geq 1\}$ is the maximal two-sided ideal in $\mathcal{O}_\pi$ (compare [48, §6.4]). Define Γ to be the image of $\mathcal{O}_D^1 \cap \left(1 + \mathcal{Q}^3\right)$ in $\mathrm{PSL}_2(\mathbb{C})$, and let N be the associated hyperbolic orbifold $\Gamma\backslash\mathbb{H}^3$.

We claim that Γ is torsion-free and hence N is a manifold. First note that $\mathcal{Q}^n = \{u \in D_\pi \mid w(u) \geq n\}$. Now for $\gamma \in \Gamma$, we have $\gamma = 1 + q$ for $q \in \mathcal{Q}^3$; from $n(\gamma) = 1$ we get that $\mathrm{tr}(\gamma) - 2 = -n(q)$ and thus $\mathrm{tr}(\gamma) - 2 \in \pi^3$ since $w(q) \geq 3$. If γ has finite order, then $\mathrm{tr}(\gamma) = \xi + \xi^{-1}$ where ξ is a root of unity. Since $\mathrm{tr}(\gamma) \in \mathcal{O}_K$, it would have to be one of $\{-1, 0, 1\}$ and none of those are 2 mod π^3. So N is a manifold.

B.2. Topological description. Let M be the hyperbolic 3-manifold $m007(3,2)$ from the Hodgson-Weeks census [39]; alternatively, M is the $(-9/2, -3/2)$ Dehn surgery on the Whitehead link L, where $+1$ surgery on L yields the figure-8 knot rather than the trefoil. Then $\mathrm{vol}(M) \approx 1.58316666$ and $H_1(M;\mathbb{Z}) = \mathbb{Z}/3\mathbb{Z} \oplus \mathbb{Z}/9\mathbb{Z}$. Let N' be the regular cover of M corresponding to any epimorphism $\pi_1(M) \to (\mathbb{Z}/3\mathbb{Z})^2$; thus $\mathrm{vol}(N') \approx 14.24849994$. We will show:

Proposition B.2. *The hyperbolic manifolds N and N' are isometric.*

Proof. We give a detailed outline, but many steps are best checked by rigorous computation; complete Sage [62] source code for this is available at [26]. From a triangulation for the alternate topological description of M as $m036(3,-1)$, Snap [29, 22] gives the group presentation

$$\pi_1(M) = \left\langle a, b \;\middle|\; aaBaabbAbb = 1,\ abbAbAAbAbb = 1 \right\rangle \tag{8}$$

where $A = a^{-1}$ and $B = b^{-1}$. Moreover, Snap rigorously checks that M is hyperbolic and that the holonomy representation $\pi_1(M) \to \mathrm{PSL}_2(\mathbb{C})$ lifts to $\rho \colon \pi_1(M) \to \mathrm{SL}_2(\mathbb{C})$ which is characterized (up to conjugacy) by $\mathrm{tr}\left(\rho(a)\right) = \mathrm{tr}\left(\rho(b)\right) = \alpha^2 + 1$ and $\mathrm{tr}\left(\rho(ab)\right) = \alpha$.

An $\mathcal{O}_K$ basis for $\mathcal{O}_D$ can be taken to be $\{1, i, x, y\}$ where $x = (i+j)/2$ and $y = (3\pi + 3\pi^2 i + \pi^2 j + \pi k)/6$. If we define

$$\overline{a} = 1 + \alpha i + \alpha x + (\alpha - 1)y \quad \text{and} \quad \overline{b} = -i \cdot \overline{a} \cdot i \tag{9}$$

then computing the norms and traces of $\left\{\overline{a},\, \overline{b},\, \overline{a}{\cdot}\overline{b}\right\}$ and evaluating the relations in (8) shows that $a \mapsto \overline{a}$ and $b \mapsto \overline{b}$ gives a homomorphism $\pi_1(M) \to \mathcal{O}_D^1 \leq \mathrm{SL}_2(\mathbb{C})$ which is a conjugate of ρ. Henceforth, we identify $\pi_1(M)$ with the subgroup of $\mathcal{O}_D^1$ generated by $\left\{\overline{a}, \overline{b}\right\}$.

Now, GAP or Magma [28, 8] easily checks that $\pi_1(N')$ is generated by

$$\left\{c = a^3, d = b^3, e = baBA, f = bABa\right\}$$

with defining relators:

$$\begin{array}{ll} DefDeceFdFcFe & DeceDecDCEfCEfCfDf \\ ECEdFcDfDeceDeccFec & fCfDecdcFecfDeceDec \end{array}$$

To see that $\pi_1(N') \leq \Gamma$, one just checks that $w(g-1) = 3$ for g in $\{c, d, e, f\}$ to confirm that each is in $1 + \mathcal{Q}^3$. By the volume formula [48, Thm 11.1.3], $\mathrm{vol}(\Lambda \backslash \mathbb{H}^3) \approx 0.26386111$ and hence $[\Lambda : \pi_1(N')] = 54$. On the other hand, one can calculate $[\Lambda : \Gamma]$ exactly as in the proof of Theorem 1.4 of [18]; while the number field in that example is $\mathbb{Q}(\sqrt{-2})$, in both examples $K_\pi \cong \mathbb{Q}_3$ and hence have isomorphic D_π. Since it turns out that $[\Lambda : \Gamma]$ is also 54, we have $\pi_1(N') = \Gamma$ as claimed. □

Theorem B.1 follows immediately from the following two lemmas, whose proofs are independent of one another.

Lemma B.3. *Let N be the closed hyperbolic 3-manifold defined above. Then N has systole $\approx 1.80203613 > 2\log\left(1+\sqrt{2}\right)$. In particular, $\pi_1(N)$ is diffuse.*

Lemma B.4. *Let N be the closed hyperbolic 3-manifold defined above. Then $\pi_1(N)$ is not left-orderable.*

Proof of Lemma B.3. We will show that the shortest geodesics in N correspond to elements $\gamma \in \Gamma = \pi_1(N)$ with $\mathrm{tr}(\gamma) = \alpha^2 - \alpha$; one such element is $ec = baBaa$. Since the translation length of γ is given by

$$\min(\gamma) = T(\mathrm{tr}(\gamma)) \quad \text{where } T(z) = \mathrm{Re}\big(2\,\mathrm{arcosh}(z/2)\big) \tag{10}$$

the systole will thus have length ≈ 1.8020361. The conclusion that Γ is diffuse follows immediately from Bowditch's criterion (iv) quoted above in Section 2.1.
We will use the Minkowski geometry of numbers picture (see e.g. [50, §I.5]) to determine the possible traces of elements of Γ with short translation lengths. Let $\tau_{\mathbb{C}}\colon K \to \mathbb{C}$ be the preferred complex embedding and $\tau_{\mathbb{R}}\colon K \to \mathbb{R}$ be the real embedding. We have the usual embedding from K into the Minkowski space $K_{\mathbb{R}} = \mathbb{R} \times \mathbb{C}$ given by $\iota = \tau_{\mathbb{R}} \times \tau_{\mathbb{C}}$, and the key fact is that $\iota(\mathcal{O}_K)$ is a lattice in $K_{\mathbb{R}}$. Thus the following set is finite:

$$\mathcal{T} = \left\{ t \in \mathcal{O}_K \;\middle|\; |\tau_{\mathbb{R}}(t)| \leq 2,\ |\tau_{\mathbb{C}}(t)| \leq 4, \text{ and } t - 2 \equiv 0 \bmod \pi^3 \right\}$$

We next show that $\mathcal{T}$ contains $\mathrm{tr}(\gamma)$ for any $\gamma \in \Gamma$ with $\min(\gamma) \leq 2.5$. That $|\tau_{\mathbb{R}}(\mathrm{tr}\,\gamma)| \leq 2$ follows since Γ is arithmetic: the quaternion algebra D ramifies at the real place and so D^1 becomes SU_2 there. To see that $\min(\gamma) \leq 2.5$ implies $|\tau_{\mathbb{C}}(t)| \leq 4$, note that $T(z)$ is minimized for fixed $|z|$ on the real axis and that $T(4) < 2.6339$.
To complete the proof of the lemma, we will show that $\mathcal{T} = \{2, \alpha^2 - \alpha, -2\alpha^2 + \alpha - 1\}$, which suffices since $T(-2\alpha^2 + \alpha - 1) \approx 2.33248166$. The natural inner product on $K_{\mathbb{R}}$ is such that $|\iota(k)|^2 = |\tau_{\mathbb{R}}(k)|^2 + 2|\tau_{\mathbb{C}}(k)|^2$ for all $k \in K$. Hence any element of $\mathcal{T}$ has norm ≤ 6, and our strategy is to enumerate all elements of $\iota(\mathcal{O}_K)$ to that norm and check which are in $\mathcal{T}$. A $\mathbb{Z}$-basis for $\mathcal{O}_K$ is $\{1, \alpha, \alpha^2\}$, and the Gram matrix in that basis for the inner product on $K_{\mathbb{R}}$ has smallest eigenvalue ≈ 1.534033. Regarding $\mathbb{Z}^3$ as having the standard norm from $\mathbb{R}^3$, this says that the natural map $\mathbb{Z}^3 \to \iota(\mathcal{O}_K)$ is distance nondecreasing. Hence every element of $\mathcal{T}$ has the form $c_0 + c_1\alpha + c_2 a^2$ where $c_i \in \mathbb{Z}$ with $c_1^2 + c_2^2 + c_3^2 \leq 36$. Computing $\mathcal{T}$ is now a simple enumeration of the 925 such triples (c_1, c_2, c_3). See [26] for a short program which does this. □

Turning to the proof of Lemma B.4, you will quickly see that it was discovered by computer, using the method of [16, §8]. Verifying its correctness is a matter of checking that 23 different elements in Γ are the identity, which can be easily done using the explicit quaternions given in (9); sample code for this is provided with [26].

Proof of Lemma B.4. Assume Γ is left-orderable and consider the positive cone $P = \{\gamma \in \Gamma \mid \gamma > 1\}$. We define some additional elements of Γ by

$$g = aBABB \quad h = abbAb \quad n = aBBAB \quad m = aBaab \quad v = ABAAb$$

By symmetry, we can assume $g \in P$. We now try all the possibilities for whether the elements $\{g, h, n, d, c, m, v\}$ are in P or not, in each case leading to the contradiction that $1 \in P$.

- Case $h \in P$:
  - Case $n \in P$:
    - Case $d \in P$:
      - Case $c \in P$:
        - Case $m \in P$: Then P contains the following, which is 1 in Γ:
          $cgndhmgmcmdhmchm$
        - Case $M \in P$: Then P contains the following, which is 1 in Γ:
          $MgndhdMgndhdMgnMhdMndMdMgndhdMgnMhdMgn$
      - Case $C \in P$: Then P contains the following, which is 1 in Γ:
        $hCggnhCgChCgd$
    - Case $D \in P$:
      - Case $c \in P$:
        - Case $m \in P$: Then P contains the following, which is 1 in Γ:
          $mcDnDmcDnDmcmnDmhDmDmcDnDmcmnDmc$
        - Case $M \in P$: Then P contains the following, which is 1 in Γ:
          $gnMDnMgnnMgncDnMg$
      - Case $C \in P$: Then P contains the following, which is 1 in Γ:
        $ChDnhCgDnhCggnnhCggnhCg$
  - Case $N \in P$:
    - Case $d \in P$: Then P contains the following, which is 1 in Γ:
      $NdhNgdNdhhNgdNdhdNdhNgdNgdNdhdNdhNgdN$
    - Case $D \in P$:
      - Case $c \in P$:
        - Case $m \in P$: Then P contains the following, which is 1 in Γ:
          $DmchNgmgmDNm$
        - Case $M \in P$:
          - Case $f \in P$: Then P contains the following, which is 1 in Γ:
            $hNcNfhMhNcNfhNhNfDf$
          - Case $F \in P$: Then P contains the following, which is 1 in Γ:
            $NhNcFMhNhNcFMhNcMhNcFcF$
      - Case $C \in P$:
        - Case $v \in P$:
          - Case $f \in P$: Then P contains the following, which is 1 in Γ:

$ChCgChNfhf$

Case $F \in P$: Then P contains the following, which is 1 in Γ:

$CvFhCgvhCgvFhCgCh$

Case $V \in P$: Then P contains the following, which is 1 in Γ:

$ChCgChNhCgChVChChNhCgChhChNhCgChVChVhhChNhCgChVChV$

Case $H \in P$:

Case $n \in P$:

Case $d \in P$: Then P contains the following, which is 1 in Γ:

$nHggdHnHgnnHggnnHggdHnHgndgnnHggnnHggdHnHgnn$

Case $D \in P$:

Case $c \in P$: Then P contains the following, which is 1 in Γ:

$DnHcHDHgnnHcHDHDnnHcHDHDnDH$

Case $C \in P$: Then P contains the following, which is 1 in Γ:

$DHCnHgnnHggnnHgnHCnHgnnCnHCnHgnn$

Case $N \in P$:

Case $d \in P$:

Case $c \in P$: Then P contains the following, which is 1 in Γ:

$NNgdHcNgdHggdHgdNNgdHcNgdHggNgdHcNgdNNgdHcd$

Case $C \in P$: Then P contains the following, which is 1 in Γ:

$NgdHggdCgdNNgdCgNgdHggdCgdHggdCgdNd$

Case $D \in P$:

Case $c \in P$:

Case $m \in P$: Then P contains the following, which is 1 in Γ:

$NcHcHDNmDHDcHcHDNmHcHDNcHccHD$

Case $M \in P$:

Case $v \in P$: Then P contains the following, which is 1 in Γ:

$HcHDMvHcHDNcHHcHDNcHcv$

Case $V \in P$: Then P contains the following, which is 1 in Γ:

$DcVcHDNcHcHDcVcHDHVcHDH$

Case $C \in P$:

Case $m \in P$: Then P contains the following, which is 1 in Γ:

$DmgmDNmDNgmgmDNmDHDDmgmDNmDHHgmDNm$

Case $M \in P$: Then P contains the following, which is 1 in Γ:

$HDHCMHCMHDMCDHCMHDMHDMCMHCMHDM$

□

References

[1] Ian Agol, *Tameness of hyperbolic 3-manifolds*, ArXiv Mathematics e-prints (2004).

[2] ———, *The virtual Haken conjecture*, Doc. Math. 18 (2013), 1045–1087, With an appendix by Agol, Daniel Groves, and Jason Manning.

[3] G. Arzhantseva and M. Steenbock, *Rips construction without unique product*, ArXiv e-prints (2014).

[4] L. Auslander and M. Kuranishi, *On the holonomy group of locally Euclidean spaces*, Ann. of Math. (2) 65 (1957), 411–415.

[5] Arthur Bartels, Wolfgang Lück, and Holger Reich, *On the Farrell-Jones conjecture and its applications*, J. Topol. 1 (2008), no. 1, 57–86.

[6] George M. Bergman, *Right orderable groups that are not locally indicable*, Pacific J. Math. 147 (1991), no. 2, 243–248.

[7] Károly J. Böröczky, Péter P. Pálfy, and Oriol Serra, *On the cardinality of sumsets in torsion-free groups*, Bull. Lond. Math. Soc. 44 (2012), no. 5, 1034–1041.

[8] Wieb Bosma, John Cannon, and Catherine Playoust, *The Magma algebra system. I. The user language*, J. Symbolic Comput. 24 (1997), no. 3-4, 235–265, Computational algebra and number theory (London, 1993).

[9] Brian H. Bowditch, *Geometrical finiteness with variable negative curvature*, Duke Math. J. 77 (1995), no. 1, 229–274.

[10] ———, *A variation on the unique product property*, J. London Math. Soc. (2) 62 (2000), no. 3, 813–826.

[11] ———, *A course on geometric group theory*, MSJ Memoirs, vol. 16, Mathematical Society of Japan, Tokyo, 2006.

[12] Steven Boyer, Dale Rolfsen, and Bert Wiest, *Orderable 3-manifold groups*, Ann. Inst. Fourier (Grenoble) 55 (2005), no. 1, 243–288.

[13] Harold Brown, Rolf Bülow, Joachim Neubüser, Hans Wondratschek, and Hans Zassenhaus, *Crystallographic groups of four-dimensional space*, John Wiley & Son, New York, 1978, Wiley Monographs in Crystallography.

[14] W. Burnside, *Notes on the theory of groups of finite order.*, Proc. Lond. Math. Soc. 26 (1895), 191–214.

[15] Peter Buser and Hermann Karcher, *Gromov's almost flat manifolds*, Astérisque, vol. 81, Société Mathématique de France, Paris, 1981.

[16] Danny Calegari and Nathan M. Dunfield, *Laminations and groups of homeomorphisms of the circle*, Invent. Math. 152 (2003), no. 1, 149–204.

[17] Danny Calegari and David Gabai, *Shrinkwrapping and the taming of hyperbolic 3-manifolds*, J. Amer. Math. Soc. 19 (2006), no. 2, 385–446.

[18] Frank Calegari and Nathan M. Dunfield, *Automorphic forms and rational homology 3-spheres*, Geom. Topol. 10 (2006), 295–329.

[19] William Carter, *New examples of torsion-free non-unique product groups*, J. Group Theory 17 (2014), no. 3, 445–464.

[20] Ted Chinburg, Eduardo Friedman, Kerry N. Jones, and Alan W. Reid, *The arithmetic hyperbolic 3-manifold of smallest volume*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 30 (2001), no. 1, 1–40.

[21] Gerald Cliff and Alfred Weiss, *Torsion free space groups and permutation lattices for finite groups*, Representation theory, group rings, and coding theory, Contemp. Math., vol. 93, Amer. Math. Soc., Providence, RI, 1989, pp. 123–132.

[22] David Coulson, Oliver A. Goodman, Craig D. Hodgson, and Walter D. Neumann, *Computing arithmetic invariants of 3-manifolds*, Experiment. Math. 9 (2000), no. 1, 127–152.

[23] David A. Craven and Peter Pappas, *On the unit conjecture for supersoluble group algebras*, J. Algebra 394 (2013), 310–356.

[24] Thomas Delzant, *Sur l'anneau d'un groupe hyperbolique*, C. R. Acad. Sci. Paris Sér. I Math. 324 (1997), no. 4, 381–384.

[25] Bertrand Deroin, Andrés Navas, and Cristóbal Rivas, *Groups, Orders, Dynamics*, 2015, Book preprint http://www.math.ens.fr/~deroin/Publications/GOD.pdf.

[26] Nathan M. Dunfield, Steffen Kionke, and Jean Raimbault, *Ancillary files with the arXiv version of this paper*, available at http://arxiv.org/abs/1411.6449.

[27] Graham Everest and Thomas Ward, *Heights of polynomials and entropy in algebraic dynamics*, Universitext, Springer-Verlag London, Ltd., London, 1999.

[28] The GAP Group, *GAP – Groups, Algorithms, and Programming, Version 4.7.5*, 2014.

[29] Oliver A. Goodman and Walter D. Neumann, *Snap for hyperbolic 3-manifolds (computer program).*

[30] M. Gromov, *Almost flat manifolds*, J. Differential Geom. 13 (1978), no. 2, 231–241.

[31] D. Gruber, A. Martin, and M. Steenbock, *Finite index subgroups without unique product in graphical small cancellation groups*, Bull. Lond. Math. Soc. 47 (2015), no. 4, 631–638.

[32] K. W. Gruenberg, *Residual properties of infinite soluble groups*, Proc. London Math. Soc. (3) 7 (1957), 29–62.

[33] Ku Yong Ha and Jong Bum Lee, *Crystallographic groups of Sol*, Math. Nachr. 286 (2013), no. 16, 1614–1667.

[34] Marshall Hall, Jr., *The theory of groups*, The Macmillan Co., New York, N.Y., 1959.

[35] Allen Hatcher, *Notes on basic three–manifold topology*, 2007, unpublished notes, available at http://www.math.cornell.edu/~hatcher/3M/3Mdownloads.html.

[36] John Hempel, 3-*Manifolds*, Princeton University Press, Princeton, N. J.; University of Tokyo Press, Tokyo, 1976, Ann. of Math. Studies, No. 86.

[37] ———, *Residual finiteness for* 3*-manifolds*, Combinatorial group theory and topology (Alta, Utah, 1984), Ann. of Math. Stud., vol. 111, Princeton Univ. Press, Princeton, NJ, 1987, pp. 379–396.

[38] Howard Hiller and Chih-Han Sah, *Holonomy of flat manifolds with* $b_1 = 0$, Quart. J. Math. Oxford Ser. (2) 37 (1986), no. 146, 177–187.

[39] Craig D. Hodgson and Jeffrey R. Weeks, *Symmetries, isometries and length spectra of closed hyperbolic three-manifolds*, Experiment. Math. 3 (1994), no. 4, 261–274.

[40] James Howie, *On locally indicable groups*, Math. Z. 180 (1982), no. 4, 445–461.

[41] William Jaco, *Lectures on three-manifold topology*, CBMS Regional Conference Series in Mathematics, vol. 43, American Mathematical Society, Providence, R.I., 1980.

[42] Kate Juschenko and Nicolas Monod, *Cantor systems, piecewise translations and simple amenable groups*, Ann. of Math. (2) 178 (2013), no. 2, 775–787.

[43] Inkang Kim and John R. Parker, *Geometry of quaternionic hyperbolic manifolds*, Math. Proc. Cambridge Philos. Soc. 135 (2003), no. 2, 291–320.

[44] Bruce Kleiner and John Lott, *Notes on Perelman's papers*, Geom. Topol. 12 (2008), no. 5, 2587–2855.

[45] P. H. Kropholler, P. A. Linnell, and J. A. Moody, *Applications of a new K-theoretic theorem to soluble group rings*, Proc. Amer. Math. Soc. 104 (1988), no. 3, 675–684.

[46] Peter Linnell and Dave Witte Morris, *Amenable groups with a locally invariant order are locally indicable*, Groups Geom. Dyn. 8 (2014), no. 2, 467–478.

[47] Wolfgang Lück, *L^2-invariants: theory and applications to geometry and K-theory*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge., vol. 44, Springer-Verlag, Berlin, 2002.

[48] Colin Maclachlan and Alan W. Reid, *The arithmetic of hyperbolic 3-manifolds*, Graduate Texts in Mathematics, vol. 219, Springer-Verlag, New York, 2003.

[49] Katsuhiko Matsuzaki and Masahiko Taniguchi, *Hyperbolic manifolds and Kleinian groups*, Oxford Mathematical Monographs, The Clarendon Press, Oxford University Press, New York, 1998, Oxford Science Publications.

[50] Jürgen Neukirch, *Algebraic number theory*, Grundlehren der Mathematischen Wissenschaften, vol. 322, Springer-Verlag, Berlin, 1999.

[51] G. Perelman, *The entropy formula for the Ricci flow and its geometric applications*, ArXiv Mathematics e-prints (2002).

[52] ———, *Ricci flow with surgery on three-manifolds*, ArXiv Mathematics e-prints (2003).

[53] Mark B. Phillips, *Dirichlet polyhedra for cyclic groups in complex hyperbolic space*, Proc. Amer. Math. Soc. 115 (1992), no. 1, 221–228.

[54] S. David Promislow, *A simple example of a torsion-free, nonunique product group*, Bull. London Math. Soc. 20 (1988), no. 4, 302–304.

[55] M. S. Raghunathan, *Discrete subgroups of Lie groups*, Springer-Verlag, New York-Heidelberg, 1972, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 68.

[56] John G. Ratcliffe, *Foundations of hyperbolic manifolds*, second ed., Graduate Texts in Mathematics, vol. 149, Springer, New York, 2006.

[57] Eliyahu Rips and Yoav Segev, *Torsion-free group without unique product property*, J. Algebra 108 (1987), no. 1, 116–126.

[58] Juan P. Rossetti and Andrzej Szczepański, *Generalized Hantzsche-Wendt flat manifolds*, Rev. Mat. Iberoamericana 21 (2005), no. 3, 1053–1070.

[59] Peter Scott, *The geometries of 3-manifolds*, Bull. London Math. Soc. 15 (1983), no. 5, 401–487.
[60] Jean-Pierre Serre, *Trees*, Springer-Verlag, Berlin, 1980.
[61] Markus Steenbock, *Rips-Segev torsion-free groups without the unique product property*, J. Algebra 438 (2015), 337–378.
[62] W. A. Stein et al., *Sage Mathematics Software (Version 5.12)*, The Sage Development Team, 2013, `http://www.sagemath.org`.
[63] Andrzej Szczepański, *Aspherical manifolds with the* $\mathbf{Q}$*-homology of a sphere*, Mathematika 30 (1983), no. 2, 291–294 (1984).
[64] William P. Thurston, *Three-dimensional manifolds, Kleinian groups and hyperbolic geometry*, Bull. Amer. Math. Soc. (N.S.) 6 (1982), no. 3, 357–381.

Steffen Kionke
Heinrich-Heine-Universität Düsseldorf
Mathematisches Institut
Universitätsstr. 1
40225 Düsseldorf
Germany
steffen.kionke@uni-duesseldorf.de

Jean Raimbault
Institut de Mathématiques de Toulouse ; UMR5219
Université de Toulouse ; CNRS
UPS IMT
F-31062 Toulouse Cedex 9
France
Jean.Raimbault@math.univ-toulouse.fr

Nathan M. Dunfield
Department of Mathematics
MC-382
University of Illinois
1409 West Green Street
Urbana
IL 61801
United States
nathan@dunfield.info

Outer Automorphisms of Algebraic Groups and a Skolem-Noether Theorem for Albert Algebras

Dem Andenken Reinhard Börgers gewidmet

Skip Garibaldi and Holger P. Petersson

Received: June 1, 2015
Revised: March 7, 2016

Communicated by Ulf Rehmann

Abstract. The question of existence of outer automorphisms of a simple algebraic group G arises naturally both when working with the Galois cohomology of G and as an example of the algebro-geometric problem of determining which connected components of $\mathrm{Aut}(G)$ have rational points. The existence question remains open only for four types of groups, and we settle one of the remaining cases, type ${}^3D_4$. The key to the proof is a Skolem-Noether theorem for cubic étale subalgebras of Albert algebras which is of independent interest. Necessary and sufficient conditions for a simply connected group of outer type A to admit outer automorphisms of order 2 are also given.

2000 Mathematics Subject Classification: Primary 20G41; Secondary 11E72, 17C40, 20G15

Contents

1 Introduction

An algebraic group H defined over an algebraically closed field F is a disjoint union of connected components. The component H° containing the identity element is a normal subgroup in H that acts via multiplication on each of the other components. Every F-point x in a connected component X of H gives an isomorphism of varieties with an H°-action $H^\circ \xrightarrow{\sim} X$ via $h \mapsto hx$.
When F is not assumed to be algebraically closed, the identity component H° is still defined as an F-subgroup of H, but the other components need not be. Suppose X is a connected subvariety of H such that, after base change to the algebraic closure F_{alg} of F, $X \times F_{\mathrm{alg}}$ is a connected component of $H \times F_{\mathrm{alg}}$. Then, by the previous paragraph, X is an H°-torsor, but X may have no F-points. We remark that the question of whether X has an F-point arises when describing the embedding of the category of compact real Lie groups into the category of linear algebraic groups over $\mathbb{R}$, see [Se, §5].

1.1 Outer automorphisms of algebraic groups

We will focus on the case where $H = \mathrm{Aut}(G)$ and G is semisimple, which amounts to asking about the existence of outer automorphisms of G. This question has previously been studied in [MT], [PrT], [Gar 12], [CKT], [CEKT], and [KT]. Writing Δ for the Dynkin diagram of G endowed with the natural action by the Galois group $\mathrm{Gal}(F_{\mathrm{sep}}/F)$ gives an exact sequence of group schemes

$$1 \longrightarrow \mathrm{Aut}(G)^\circ \longrightarrow \mathrm{Aut}(G) \xrightarrow{\alpha} \mathrm{Aut}(\Delta)$$

as in [DG, Chap. XXIV, Th. 1.3 and §3.6] or [Sp, §16.3], hence a natural map $\alpha(F) : \mathrm{Aut}(G)(F) \to \mathrm{Aut}(\Delta)(F)$. Note that $\mathrm{Aut}(\Delta)(F_{\mathrm{alg}})$ is identified with the connected components of $\mathrm{Aut}(G) \times F_{\mathrm{alg}}$ in such a way that $\mathrm{Aut}(\Delta)(F)$ is identified with those components that are defined over F. We ask: is $\alpha(F)$ onto? That is, *which of the components of* $\mathrm{Aut}(G)$ *that are defined over F also have an F-point?*
Sending an element g of G to conjugation by g defines a surjection $G \to \mathrm{Aut}(G)^\circ$, and the F-points $\mathrm{Aut}(G)^\circ(F)$ are called *inner* automorphisms. The F-points of the other components of $\mathrm{Aut}(G)$ are called *outer*. Therefore, our question may be rephrased as: *Is every automorphism of the Dynkin diagram that is compatible with the natural action by the Galois group of F induced from an F-automorphism of G?*

One can quickly observe that $\alpha(F)$ need not be onto, for example, with the group $\mathrm{SL}(A)$ where A is a central simple algebra of odd exponent, where an outer automorphism would amount to an isomorphism of A with its opposite algebra. This is a special case of a general cohomological obstruction. Namely, writing Z for the scheme-theoretic center of the simply connected cover of G, G naturally defines an element $t_G \in H^2(F, Z)$ called the *Tits class* as in [T, 4.2] or [KMRT, 31.6]. (The cohomology used in this paper is fppf.) For every character $\chi\colon Z \to \mathbb{G}_{\mathrm{m}}$, the image $\chi(t_G) \in H^2(F, \mathbb{G}_{\mathrm{m}})$ is known as a Tits algebra of G; for example, when $G = \mathrm{SL}(A)$, Z is identified with the group of $(\deg A)$-th roots of unity, the group of characters is generated by the natural inclusion $\chi\colon Z \hookrightarrow \mathbb{G}_{\mathrm{m}}$, and $\chi(t_{\mathrm{SL}(A)})$ is the class of A. (More such examples are given in [KMRT, §27.B].) This example illustrates also the general fact: $t_G = 0$ if and only if $\mathrm{End}_G(V)$ is a field for every irreducible representation V of G. The group scheme $\mathrm{Aut}(\Delta)$ acts on $H^2(F, Z)$, and it was shown in [Gar 12, Th. 11] that this provides an obstruction to the surjectivity of $\alpha(F)$, namely:

$$\mathrm{im}\,[\alpha\colon \mathrm{Aut}(G)(F) \to \mathrm{Aut}(\Delta)(F)] \subseteq \{\pi \in \mathrm{Aut}(\Delta)(F) \mid \pi(t_G) = t_G\}. \quad (1.1.1)$$

It is interesting to know when equality holds in (1.1.1), because this information is useful in Galois cohomology computations. (For example, when G is simply connected, equality in (1.1.1) is equivalent to the exactness of $H^1(F, Z) \to H^1(F, G) \to H^1(F, \mathrm{Aut}(G))$.) Certainly, equality need not hold in (1.1.1), for example when G is semisimple (take G to be the product of the compact and split real forms of G_2) or when G is neither simply connected nor adjoint (take G to be the split group SO_8, for which $|\mathrm{im}\,\alpha| = 2$ but the right side of (1.1.1) has 6 elements). However, when G is simple and simply connected or adjoint, it is known that equality holds in (1.1.1) when G has inner type or for some fields F. Therefore, one might optimistically hope that the following is true:

Conjecture 1.1.2. *If G is an absolutely simple algebraic group that is simply connected or adjoint, then equality holds in* (1.1.1).

The remaining open cases are where G has type ${}^2A_n$ for odd $n \geq 3$ (the case where n is even is Cor. 9.1.2), ${}^2D_n$ for $n \geq 3$, ${}^3D_4$, and ${}^2E_6$. Most of this paper is dedicated to settling one of these four cases.

Theorem A. *If G is a simple algebraic group of type ${}^3D_4$ over a field F, then equality holds in* (1.1.1).

One can ask also for a stronger property to hold:

Question 1.1.3. *Suppose π is in $\alpha(\mathrm{Aut}(G)(F))$. Does there exist a $\phi \in \mathrm{Aut}(G)(F)$ so that $\alpha(\phi) = \pi$ and ϕ and π have the same order?*

This question, and a refinement of it where one asks for detailed information about the possible ϕ's, was considered for example in [MT], [PrT], [CKT], [CEKT], and [KT]. (The paper [Br] considers a different but related question, on the level of group schemes and not k-points.) It was observed in [Gar 12] that the answer to Question 1.1.3 is "yes" in all the cases where Conjecture 1.1.2 is known to hold. However, [KT] gives an example of a group G of type ${}^3D_4$ that does not have an outer automorphism of order 3, yet the conjecture holds for G by Theorem A. That is, combining the results of this paper and [KT] gives the first example where the conjecture holds for a group but the answer to Question 1.1.3 is "no", see Example 8.3.1.

In the final section of the paper, §9, we translate the conjecture for groups of type A into one in the language of algebras with involution as in [KMRT], give a criterion for the existence of outer automorphisms of order 2 (i.e., prove a version for type A of the main result of [KT]), and exhibit a group of type ${}^2A$ that does not have an outer automorphism of order 2.

1.2 Skolem-Noether Theorem for Albert algebras

In order to prove Theorem A, we translate it into a statement about Albert F-algebras, 27-dimensional exceptional central simple Jordan algebras. Specifically, we realize a simply connected group G of type ${}^3D_4$ with $t_G = 0$ as a subgroup of the structure group of an Albert algebra J that fixes a cyclic cubic subfield E elementwise, as in [KMRT, 38.7]. For such a group, the right side of (1.1.1) is $\mathbb{Z}/3$ and we prove equality in (1.1.1) by extending, in a controlled way, a nontrivial F-automorphism of E to J, see the proof of Prop. 8.2.2. The desired extension exists by Theorem B below, whose proof is the focus of §§2–7. We spend the majority of the paper working with Jordan algebras.

Let J be an Albert algebra over a field F and suppose $E, E' \subseteq J$ are cubic étale subalgebras. It is known since Albert-Jacobson [AJ] that in general an isomorphism $\varphi\colon E \to E'$ cannot be extended to an automorphism of J. Thus the Skolem-Noether Theorem fails to hold for cubic étale subalgebras of Albert algebras. In fact, even in the important special case that $E = E'$ is split and φ is an automorphism of E having order 3, obstructions to the validity of this result may be read off from [AJ, Th. 9]. We provide a way out of this impasse by replacing the automorphism group of J by its structure group and allowing the isomorphism φ to be twisted by the right multiplication of a norm-one element in E. More precisely, referring to our notational conventions in Sections 1.3−3 below, we will establish the following result. For $w \in E$, write $R_w\colon E \to E$ for the right multiplication $e \mapsto ew$.

Theorem B. *Let $\varphi\colon E \xrightarrow{\sim} E'$ be an isomorphism of cubic étale subalgebras of an Albert algebra J over a field F. Then there exists an element $w \in E$ satisfying $N_E(w) = 1$ such that $\varphi \circ R_w\colon E \to E'$ can be extended to an element of the structure group of J.*

Note that no restrictions on the characteristic of F will be imposed. In order to prove Theorem B, we first derive its analogue (in fact, a substantial generalization of it, see Th. 5.2.7 below) for absolutely simple Jordan algebras of degree 3 and dimension 9 in place of J. This generalization is based on the notions of weak and strong equivalence for isotopic embeddings of cubic étale algebras into cubic Jordan algebras (4.1) and is derived here by elementary manipulations of the two Tits constructions. After a short digression into norm classes for pairs of isotopic embeddings in § 6, Theorem B is established by combining Th. 5.2.7 with a density argument and the fact that an isotopy between absolutely simple nine-dimensional subalgebras of an Albert algebra can always be extended to an element of its structure group (Prop. 7.2.4).

1.3 Conventions.

Throughout this paper, we fix a base field F of arbitrary characteristic. All linear non-associative algebras over F are tacitly assumed to contain an identity element. If C is such an algebra, we write $C^\times$ for the collection of invertible elements in C, whenever this makes sense. For any commutative associative algebra K over F, we denote by $C_K := C \otimes K$ the scalar extension (or base change) of C from F to K, unadorned tensor products always being taken over F. In other terminological and notational conventions, we mostly follow [KMRT]. In fact, the sole truly significant deviation from this rule is presented by the theory of Jordan algebras: while [KMRT, Chap. IX] confines itself to the linear version of this theory, which works well only over fields of characteristic not 2 or, more generally, over commutative rings containing $\frac{1}{2}$, we insist on the quadratic one, surviving as it does in full generality over arbitrary commutative rings. For convenience, we will assemble the necessary background material in the next two sections of this paper.

2 Jordan algebras

The purpose of this section is to present a survey of the standard vocabulary of arbitrary Jordan algebras. Our main reference is [J 81].

2.1 The concept of a Jordan algebra

By a (unital quadratic) *Jordan algebra* over F, we mean an F-vector space J together with a quadratic map $x \mapsto U_x$ from J to $\mathrm{End}_F(J)$ (the U*-operator*) and a distinguished element $1_J \in J$ (the *unit* or *identity element*) such that, writing

$$\{xyz\} := V_{x,y}z := U_{x,z}y := (U_{x+z} - U_x - U_z)y$$

for the associated *triple product*, the equations

$$\begin{aligned} U_{1_J} &= \mathbf{1}_J, \\ U_{U_x y} &= U_x U_y U_x \qquad \text{(fundamental formula)}, \\ U_x V_{y,x} &= V_{x,y} U_x \end{aligned} \tag{2.1.1}$$

hold in all scalar extensions. We always simply write J to indicate a Jordan algebra over F, U-operator and identity element being understood. A *subalgebra* of J is an F-subspace containing the identity element and stable under the operation U_xy; it is then a Jordan algebra in its own right. A *homomorphism* of Jordan algebras over F is an F-linear map preserving U-operators and identity elements. In this way we obtain the category of Jordan algebras over F. By definition, the property of being a Jordan algebra is preserved by arbitrary scalar extensions. In keeping with the conventions of Section 1.3, we write J_K for the base change of J from F to any commutative associative F-algebra K.

2.2 Linear Jordan algebras

Assume $\mathrm{char}(F) \neq 2$. Then Jordan algebras as defined in 2.1 and linear Jordan algebras as defined in [KMRT, § 37] are virtually the same. Indeed, let J be a unital quadratic Jordan algebra over F. Then J becomes an ordinary non-associative F-algebra under the multiplication $x \cdot y := \frac{1}{2}U_{x,y}1_J$, and this F-algebra is a linear Jordan algebra in the sense that it is commutative and satisfies the *Jordan identity* $x \cdot ((x \cdot x) \cdot y) = (x \cdot x) \cdot (x \cdot y)$. Conversely, let J be a linear Jordan algebra over F. Then the U-operator $U_xy := 2x \cdot (x \cdot y) - (x \cdot x) \cdot y$ and the identity element 1_J convert J into a unital quadratic Jordan algebra. The two constructions are inverse to one another and determine an isomorphism of categories between unital quadratic Jordan algebras and linear Jordan algebras over F.

2.3 Ideals and simplicity

Let J be a Jordan algebra over F. A subspace $I \subseteq J$ is said to be an *ideal* if $U_IJ + U_JI + \{IIJ\} \subseteq J$. In this case, the quotient space J/I carries canonically the structure of a Jordan algebra over F such that the projection $J \to J/I$ is a homomorphism. A Jordan algebra is said to be *simple* if it is non-zero and there are no ideals other than the trivial ones. We speak of an *absolutely simple* Jordan algebra if it stays simple under all base field extensions. (There is also a notion of central simplicity which, however, is weaker than absolute simplicity, although the two agree for $\mathrm{char}(F) \neq 2$.)

2.4 Standard examples

First, let A be an associative F-algebra. Then the vector space A together with the U-operator $U_xy := xyx$ and the identity element 1_A is a Jordan algebra over F, denoted by A^+. If A is simple, then so is A^+ [McC 69b, Th. 4]. Next, let (B, τ) be an F-algebra with involution, so B is a non-associative algebra over F and $\tau\colon B \to B$ is an F-linear anti-automorphism of period 2. Then

$$H(B, \tau) := \{x \in B \mid \tau(x) = x\}$$

is a subspace of B. Moreover, if B is associative, then $H(B, \tau)$ is a subalgebra of B^+, hence a Jordan algebra which is simple if (B, τ) is simple as an algebra with involution [McC 69b, Th. 5].

2.5 Powers

Let J be a Jordan algebra over F. The powers of $x \in J$ with integer exponents $n \geq 0$ are defined recursively by $x^0 = 1_J$, $x^1 = x$, $x^{n+2} = U_x x^n$. Note for $J = A^+$ as in 2.4, powers in J and in A are the same. For J arbitrary, they satisfy the relations

$$U_{x^m} x^n = x^{2m+n}, \quad \{x^m x^n x^p\} = 2x^{m+n+p}, \quad (x^m)^n = x^{mn}, \qquad (2.5.1)$$

hence force

$$F[x] := \sum_{n \geq 0} F x^n$$

to be a subalgebra of J. In many cases — e.g., if $\mathrm{char}(F) \neq 2$ or if J is simple (but not always [J 81, 1.31, 1.32]) — there exists a commutative associative F-algebra R, necessarily unique, such that $F[x] = R^+$ [McC 70, Prop. 1], [J 81, Prop. 4.6.2]. By abuse of language, we simply write $R = F[x]$ and say R is a subalgebra of J.

In a slightly different, but similar, vein we wish to talk about étale subalgebras of a Jordan algebra. This is justified by the fact that étale F-algebras are completely determined by their Jordan structure. More precisely, we have the following simple result.

Lemma 2.5.2. *Let E, R be commutative associative F-algebras such that E is finite-dimensional étale. Then $\varphi\colon E^+ \xrightarrow{\sim} R^+$ is an isomorphism of Jordan algebras if and only if $\varphi\colon E \xrightarrow{\sim} R$ is an isomorphism of commutative associative algebras.*

Proof. Extending scalars if necessary, we may assume that E as a (unital) F-algebra is generated by a single element $x \in E$, since this is easily seen to hold unless $F = \mathbb{F}_2$, the field with two elements. But since the powers of x in E agree with those in $E^+ = R^+$, hence with those in R, the assertion follows. □

2.6 Inverses and Jordan division algebras

Let J be a Jordan algebra over F. An element $x \in J$ is said to be *invertible* if the U-operator $U_x\colon J \to J$ is bijective (equivalently, $1_J \in \mathrm{Im}(U_x)$), in which case we call $x^{-1} := U_x^{-1} x$ the *inverse* of x in J. Invertibility and inverses are preserved by homomorphisms. It follows from the fundamental formula (2.1.1) that, if $x, y \in J$ are invertible, then so is $U_x y$ and $(U_x y)^{-1} = U_{x^{-1}} y^{-1}$. Moreover, setting $x^n := (x^{-1})^{-n}$ for $n \in \mathbb{Z}$, $n < 0$, we have (2.5.1) for all $m, n, p \in \mathbb{Z}$. In agreement with earlier conventions, the set of invertible elements in J will be denoted by $J^\times$. If $J^\times = J \setminus \{0\} \neq \emptyset$, then we call J a *Jordan division algebra*. If A is an associative algebra, then $(A^+)^\times = A^\times$, and the inverses are the same. Similarly, if (B, τ) is an associative algebra with involution, then $H(B, \tau)^\times = H(B, \tau) \cap B^\times$, and, again, the inverses are the same.

2.7 Isotopes

Let J be a Jordan algebra over F and $p \in J^\times$. Then the vector space J together with the U-operator $U_x^{(p)} := U_x U_p$ and the distinguished element $1_J^{(p)} := p^{-1}$ is a Jordan algebra over F, called the *p-isotope* (or simply an *isotope*) of J and denoted by $J^{(p)}$. We have $J^{(p)\times} = J^\times$ and $(J^{(p)})^{(q)} = J^{(U_p q)}$ for all $q \in J^\times$, which implies $(J^{(p)})^{(q)} = J$ for $q := p^{-2}$. Passing to isotopes is functorial in the following sense: if $\varphi\colon J \to J'$ is a homomorphism of Jordan algebras, then so is $\varphi\colon J^{(p)} \to J'^{(\varphi(p))}$, for any $p \in J^\times$.

Let A be an associative algebra over F and $p \in (A^+)^\times = A^\times$. Then right multiplication by p in A gives an isomorphism $R_p\colon (A^+)^{(p)} \xrightarrow{\sim} A^+$ of Jordan algebras. On the other hand, if (B, τ) is an associative algebra with involution, then so is $(B, \tau^{(p)})$, for any $p \in H(B, \tau)^\times$, where $\tau^{(p)}\colon B \to B$ via $x \mapsto p^{-1}\tau(x)p$ stands for the *p-twist* of τ, and

$$R_p\colon H(B,\tau)^{(p)} \xrightarrow{\sim} H(B, \tau^{(p)}) \tag{2.7.1}$$

is an isomorphism of Jordan algebras.

2.8 Homotopies and the structure group

If J, J' are Jordan algebras over F, a *homotopy* from J to J' is a homomorphism $\varphi\colon J \to J'^{(p')}$ of Jordan algebras, for some $p' \in J'^\times$. In this case, $p' = \varphi(1_J)^{-1}$ is uniquely determined by φ. Bijective homotopies are called *isotopies*, while injective homotopies are called *isotopic embeddings*. The set of isotopies from J to itself is a subgroup of $\mathrm{GL}(J)$, called the *structure group* of J and denoted by $\mathrm{Str}(J)$. It consists of all linear bijections $\eta\colon J \to J$ such that some linear bijection $\eta^\sharp\colon J \to J$ satisfies $U_{\eta(x)} = \eta U_x \eta^\sharp$ for all $x \in J$. The structure group contains the automorphism group of J as a subgroup; more precisely, $\mathrm{Aut}(J)$ is the stabilizer of 1_J in $\mathrm{Str}(J)$. Finally, thanks to the fundamental formula (2.1.1), we have $U_y \in \mathrm{Str}(J)$ for all $y \in J^\times$.

3 Cubic Jordan algebras

In this section, we recall the main ingredients of the approach to a particularly important class of Jordan algebras through the formalism of cubic norm structures. Our main references are [McC 69a] and [JK]. Systematic use will be made of the following notation: given a polynomial map $P\colon V \to W$ between vector spaces V, W over F and $y \in V$, we denote by $\partial_y P\colon V \to W$ the polynomial map given by the derivative of P in the direction y, so $(\partial_y P)(x)$ for $x \in V$ is the coefficient of the variable $\mathbf{t}$ in the expansion of $P(x + \mathbf{t}y)$:

$$P(x + \mathbf{t}y) = P(x) + \mathbf{t}(\partial_y P)(x) + \cdots .$$

3.1 Cubic norm structures

By a *cubic norm structure* over F we mean a quadruple $X = (V, c, \sharp, N)$ consisting of a vector space V over F, a distinguished element $c \in V$ (the *base point*), a quadratic map $x \mapsto x^\sharp$ from V to V (the *adjoint*), with bilinearization $x \times y := (x+y)^\sharp - x^\sharp - y^\sharp$, and a cubic form $N\colon V \to F$ (the *norm*), such that, writing

$$T(y,z) := (\partial_y N)(c)(\partial_z N)(c) - (\partial_y \partial_z N)(c) \qquad (y, z \in V)$$

for the (bilinear) *trace* of X and $T(y) := T(y,c)$ for the linear one, the equations

$$c^\sharp = c,\ N(c) = 1 \qquad \text{(base point identities)}, \tag{3.1.1}$$
$$c \times x = T(x)c - x \qquad \text{(unit identity)}, \tag{3.1.2}$$
$$(\partial_y N)(x) = T(x^\sharp, y) \qquad \text{(gradient identity)}, \tag{3.1.3}$$
$$x^{\sharp\sharp} = N(x)x \qquad \text{(adjoint identity)} \tag{3.1.4}$$

hold in all scalar extensions. A subspace of V is called a *cubic norm substructure* of X if it contains the base point and is stable under the adjoint map; it may then canonically be regarded as a cubic norm structure in its own right. A *homomorphism* of cubic norm structures is a linear map of the underlying vector spaces preserving base points, adjoints and norms. A cubic norm structure X as above is said to be *non-singular* if V has finite dimension over F and the bilinear trace $T\colon V \times V \to F$ is a non-degenerate symmetric bilinear form. If X and Y are cubic norm structures over F, with Y non-singular, and $\varphi\colon X \to Y$ is a surjective linear map preserving base points and norms, then φ is an isomorphism of cubic norm structures [McC 69a, p. 507].

3.2 The associated Jordan algebra

Let $X = (V, c, \sharp, N)$ be a cubic norm structure over F and write T for its bilinear trace. Then the U-operator

$$U_x y := T(x,y)x - x^\sharp \times y \tag{3.2.1}$$

and the base point c convert the vector space V into a Jordan algebra over F, denoted by $J(X)$ and called the Jordan algebra *associated with* X. The construction of $J(X)$ is clearly functorial in X. We have

$$N(U_x y) = N(x)^2 N(y) \qquad (x, y \in J). \tag{3.2.2}$$

Jordan algebras isomorphic to $J(X)$ for some cubic norm structure X over F are said to be *cubic*. For example, let J be a Jordan algebra over F that is generically algebraic (e.g., finite-dimensional) of degree 3 over F. Then $X = (V, c, \sharp, N)$, where V is the vector space underlying J, $c := 1_J$, $\sharp$ is the numerator of the inversion map, and $N := N_J$ is the generic norm of J, is a cubic norm structure over F satisfying $J = J(X)$; in particular, J is a cubic Jordan algebra. In view of this correspondence, we rarely distinguish carefully between a cubic norm structure and its associated Jordan algebra. Non-singular cubic Jordan algebras, i.e., Jordan algebras arising from non-singular cubic norm structures, by [McC 69a, p. 507] have no absolute zero divisors, so $U_x = 0$ implies $x = 0$.

3.3 Cubic étale algebras

Let E be a cubic étale F-algebra. Then Lemma 2.5.2 allows us to identify $E = E^+$ as a generically algebraic Jordan algebra of degree 3 (with U-operator $U_x y = x^2 y$), so we may write $E = E^+ = J(V, c, \sharp, N)$ as in 3.2, where $c = 1_E$ is the unit element, $\sharp$ is the adjoint and $N = N_E$ is the norm of $E = E^+$. We also write T_E for the (bilinear) trace of E. The discriminant (algebra) of E will be denoted by $\Delta(E)$; it is a quadratic étale F-algebra [KMRT, 18.C].

3.4 Isotopes of cubic norm structures

Let $X = (V, c, \sharp, N)$ be a cubic norm structure over F. An element $p \in V$ is invertible in $J(X)$ if and only if $N(p) \neq 0$, in which case $p^{-1} = N(p)^{-1} p^\sharp$. Moreover,

$$X^{(p)} := (V, c^{(p)}, \sharp^{(p)}, N^{(p)}),$$

with $c^{(p)} := p^{-1}$, $x^{\sharp^{(p)}} := N(p) U_p^{-1} x^\sharp$, $N^{(p)} := N(p)N$, is again a cubic norm structure over F , called the *p-isotope* of X. This terminology is justified since the associated Jordan algebra $J(X^{(p)}) = J(X)^{(p)}$ is the p-isotope of $J(X)$. We also note that the bilinear trace of $X^{(p)}$ is given by

$$T^{(p)}(y, z) = T(U_p y, z) \qquad (y, z \in X) \qquad (3.4.1)$$

in terms of the bilinear trace T of X. Combining the preceding considerations with 3.1, we conclude that the structure group of a *non-singular* cubic Jordan algebra agrees with its group of norm similarities.

3.5 Cubic Jordan matrix algebras

Let C be a composition algebra over F, so C is a Hurwitz algebra in the sense of [KMRT, §33C], with norm n_C, trace t_C, and conjugation $v \mapsto \bar{v} := t_C(v) 1_C - v$. Note in particular that the base field itself is a composition even if it has characteristic 2. For any diagonal matrix

$$\Gamma = \mathrm{diag}(\gamma_1, \gamma_2, \gamma_3) \in \mathrm{GL}_3(F),$$

the pair

$$(\mathrm{Mat}_3(C), \tau_\Gamma), \quad \tau_\Gamma(x) := \Gamma^{-1} \bar{x}^t \Gamma \qquad (x \in \mathrm{Mat}_3(C)),$$

is a non-associative F-algebra with involution, allowing us to consider the subspace $\mathrm{Her}_3(C,\Gamma) \subseteq \mathrm{Mat}_3(C)$ consisting of all elements $x \in \mathrm{Mat}_3(C)$ that are Γ-hermitian ($x = \Gamma^{-1}\bar{x}^t\Gamma$) and have scalars down the diagonal. Note that we have

$$\mathrm{Her}_3(C,\Gamma) \subseteq H(\mathrm{Mat}_3(C), \tau_\Gamma)$$

in the sense of 2.4, with equality for $\mathrm{char}(F) \neq 2$ but not in general. In terms of the usual matrix units $e_{ij} \in \mathrm{Mat}_3(C)$, $1 \leq i, j \leq 3$, we therefore have

$$\mathrm{Her}_3(C,\Gamma) = \sum (Fe_{ii} + C[jl]),$$

the sum on the right being taken over all cyclic permutations (ijl) of (123), where

$$C[jl] := \{v[jl] \mid v \in C\}, \quad v[jl] := \gamma_l v e_{jl} + \gamma_j \bar{v} e_{lj}.$$

Now put $V := \mathrm{Her}_3(C,\Gamma)$ as a vector space over F, $c := \mathbf{1}_3$ (the 3×3 unit matrix) and define adjoint and norm on V by

$$\begin{aligned} x^\sharp &:= \sum \Big(\big(\alpha_j\alpha_l - \gamma_j\gamma_l n_C(v_i)\big)e_{ii} + \big(-\alpha_i v_i + \gamma_i \overline{v_j v_l}\big)[jl] \Big), \\ N(x) &:= \alpha_1\alpha_2\alpha_3 - \sum \gamma_j\gamma_l\alpha_i n_C(v_i) + \gamma_1\gamma_2\gamma_3 t_C(v_1v_2v_3) \end{aligned}$$

for all $x = \sum(\alpha_i e_{ii} + v_i[jl])$ in all scalar extensions of V. Then $X := (V, c, \sharp, N)$ is a cubic norm structure over F. Henceforth, the symbol $\mathrm{Her}_3(C,\Gamma)$ will stand for this cubic norm structure but also for its associated cubic Jordan algebra. We always abbreviate $\mathrm{Her}_3(C) := \mathrm{Her}_3(C, \mathbf{1}_3)$.

3.6 Albert algebras

Writing $\mathrm{Zor}(F)$ for the split octonion algebra of Zorn vector matrices over F [KMRT, VIII, Exc. 5], the cubic Jordan matrix algebra $\mathrm{Her}_3(\mathrm{Zor}(F))$ is called the *split Albert algebra* over F. By an *Albert algebra* over F, we mean an F-form of $\mathrm{Her}_3(\mathrm{Zor}(F))$, i.e., a Jordan algebra over F (necessarily absolutely simple and non-singular of degree 3 and dimension 27) that becomes isomorphic to the split Albert algebra when extending scalars to the separable closure. Albert algebras are either *reduced*, hence have the form $\mathrm{Her}_3(C,\Gamma)$ as in 3.5, C an octonion algebra over F (necessarily unique), or are cubic Jordan division algebras.

3.7 Associative algebras of degree 3 with unitary involution

By an *associative algebra of degree* 3 *with unitary involution* over F we mean a triple (K, B, τ) with the following properties: K is a quadratic étale F-algebra, with norm n_K, trace t_K and conjugation ι_K, $a \mapsto \bar{a}$, B is an associative algebra of degree 3 over K and $\tau\colon B \to B$ is an F-linear involution that induces the conjugation of K via restriction. All this makes obvious sense even in the special case that $K \cong F \times F$ is split, as do the generic norm, trace and adjoint of B, which are written as $N_B, T_B, \sharp$, respectively, connect naturally with the involution τ and agree with the corresponding notions for the cubic Jordan algebra B^+. In particular, $H(B,\tau)$ is a Jordan algebra of degree 3 over F whose associated cubic norm structure derives from that of B^+ via restriction. Let (K,B,τ) be an associative algebra of degree 3 with unitary involution over F. We say (K,B,τ) is *non-singular* if the corresponding cubic Jordan algebra B^+ has this property, equivalently, if B is free of finite rank over K and $T_B\colon B\times B\to K$ is a non-degenerate symmetric bilinear form in the usual sense. We say (K,B,τ) is *central simple* if K is the centre of B and (B,τ) is simple as an algebra with involution. This allows us to speak of (B,τ) as a *central simple associative algebra of degree* 3 *with unitary involution* over F, the centre of B (a quadratic étale F-algebra) being understood.

3.8 The second Tits construction

Let (K,B,τ) be an associative algebra of degree 3 with unitary involution over F and suppose we are given a *norm pair* (u,μ) of (K,B,τ), i.e., a pair of invertible elements $u\in H(B,\tau)$, $\mu\in K$ such that $N_B(u)=n_K(\mu)$. We put $V:=H(B,\tau)\oplus Bj$ as the external direct sum of $H(B,\tau)$ and B as vector spaces over F, using j as a placeholder. We define base point, adjoint and norm on V by the formulas

$$c := 1_B + 0\cdot j, \tag{3.8.1}$$

$$x^\sharp := (v_0^\sharp - vu\bar{v}) + (\bar{\mu}\bar{v}^\sharp u^{-1} - v_0 v)j, \tag{3.8.2}$$

$$N(x) := N_B(v_0) + \mu N_B(v) + \bar{\mu}\overline{N_B(v)} - T_B\big(v_0, vu\tau(v)\big) \tag{3.8.3}$$

for $x = v_0 + vj$, $v_0\in H(B,\tau)$, $v\in B$ (and in all scalar extensions as well). Then we obtain a cubic norm structure $X:=(V,c,\sharp,N)$ over F whose associated cubic Jordan algebra will be denoted by $J:=J(K,B,\tau,u,\mu):=J(X)$ and has the bilinear trace

$$\begin{aligned} T(x,y) &= T_B(v_0,w_0) + T_B\big(vu\tau(w)\big) + T_B\big(wu\tau(v)\big) \\ &= T_B(v_0,w_0) + t_K\Big(T_B\big(vu\tau(w)\big)\Big) \end{aligned} \tag{3.8.4}$$

for x as above and $y = w_0 + wj$, $w_0\in H(B,\tau)$, $w\in B$. It follows that, if (K,B,τ) is non-singular, then so is J. Note also that the cubic Jordan algebra $H(B,\tau)$ identifies with a subalgebra of J through the initial summand.

If (K,B,τ) is central simple in the sense of 3.7, then K is the centre of B, $J(B,\tau,u,\mu) := J(K,B,\tau,u,\mu)$ is an Albert algebra over F, and all Albert algebras can be obtained in this way. More precisely, every Albert algebra J over F contains a subalgebra isomorphic to $H(B,\tau)$ for some central simple associative algebra (B,τ) of degree 3 with unitary involution over F, and every homomorphism $H(B,\tau)\to J$ can be extended to an isomorphism from $J(B,\tau,u,\mu)$ to J, for some norm pair (u,μ) of (K,B,τ), with K the centre of B.

Our next result is a variant of [PeR 84b, Prop. 3.9] which extends the isomorphism (2.7.1) in a natural way.

Lemma 3.8.5. *Let (K, B, τ) be a non-singular associative algebra of degree 3 with unitary involution over F and suppose (u, μ) is a norm pair of (K, B, τ). Then, given any $p \in H(B, \tau)^\times$, writing $\tau^{(p)}$ for the p-twist of τ in the sense of 2.7 and setting $u^{(p)} := p^\sharp u$, $\mu^{(p)} := N_B(p)\mu$, the following statements hold.*

a. $(K, B, \tau^{(p)})$ is a non-singular associative algebra of degree 3 with unitary involution over F.

b. $H(B, \tau^{(p)}) = H(B, \tau)p$, and $(u^{(p)}, \mu^{(p)})$ is a norm pair of $(K, B, \tau^{(p)})$.

c. The map

$$\hat{R}_p\colon\ J(K, B, \tau, u, \mu)^{(p)} \xrightarrow{\ \sim\ } J(K, B, \tau^{(p)}, u^{(p)}, \mu^{(p)})$$

defined via $v_0 + vj \longmapsto v_0 p + (p^{-1}vp)j$ is an isomorphism of cubic Jordan algebras.

Proof. (a): This follows immediately from 3.7.
(b): The first assertion is a consequence of (2.7.1). As to the second, we clearly have $u^{(p)} \in B^\times$ and $\mu^{(p)} \in K^\times$. Moreover, from $p^{-1} = N_B(p)^{-1}p^\sharp$ we deduce $N_B(p^\sharp) = N_B(p)^2$ and $pp^\sharp = N_B(p)1_B = p^\sharp p$, hence $\tau^{(p)}(u^{(p)}) = p^{-1}\tau(u)p^\sharp p = N_B(p)p^{-1}u = p^\sharp u = u^{(p)}$. Thus $u^{(p)} \in H(B, \tau^{(p)})^\times$ and $N_B(u^{(p)}) = N_B(p)^2 n_B(u) = N_B(u)^2 n_K(\mu) = n_K(\mu^{(p)})$, which completes the proof.
(c): By (b), (3.4.1) and 3.8, the map $\hat{R}_p$ is a linear bijection between non-singular cubic Jordan algebras preserving base points. By 3.1, it therefore suffices to show that it preserves norms as well. Writing N (resp. N') for the norm of $J(K, B, \tau, u, \mu)$ (resp. $J(K, B, \tau^{(p)}, u^{(p)}, \mu^{(p)})$), we let $v_0 \in H(B, \tau)$, $v \in B$ and compute, using (3.8.3),

$$\begin{aligned}
(N' \circ \hat{R}_p)(v_0 + vj) &= N'(v_0 p + (p^{-1}vp)j) \\
&= N_B(p)N_B(v_0) + N_B(p)\mu N_B(v) + N_B(p)\bar{\mu}\overline{N_B(v)} \\
&\quad - T_B\big(v_0 p p^{-1} v p p^\sharp u \tau^{(p)}(p^{-1}vp)\big) \\
&= N_B(p)\Big(N_B(v_0) + \mu N_B(v) + \bar{\mu}\overline{N_B(v)} - T_B\big(v_0 v u \tau(v)\big)\Big) \\
&= N^{(p)}(v_0 + vj),
\end{aligned}$$

as desired. □

Remark 3.8.6. The lemma holds without the non-singularity condition on (K, B, τ) but the proof is more involved.

If the quadratic étale F-algebra K in 3.8 is split, there is a less cumbersome way of describing the output of the second Tits construction.

3.9 The first Tits construction

Let A be an associative algebra of degree 3 over F and $\mu \in F^\times$. Put $V := A \oplus Aj_1 \oplus Aj_2$ as the direct sum of three copies of A as an F-vector space and define base point, adjoint and norm on V by the formulas $c := 1_A + 0 \cdot j_1 + 0 \cdot j_2$,

$$x^\sharp := (v_0^\sharp - v_1 v_2) + (\mu^{-1} v_2^\sharp - v_0 v_1) j_1 + (\mu v_1^\sharp - v_2 v_0) j_2, \tag{3.9.1}$$

$$N(x) := N_A(v_0) + \mu N_A(v_1) + \mu^{-1} N_A(v_2) - T_A(v_0 v_1 v_2) \tag{3.9.2}$$

for $x = v_0 + v_1 j_1 + v_2 j_2$, v_0, v_1, v_2 running over all scalar extensions of A. Then $X := (V, c, \sharp, N)$ is a cubic norm structure over F, with bilinear trace given by

$$T(x, y) = T_A(v_0, w_0) + T_A(v_1, w_2) + T_A(v_2, w_1) \tag{3.9.3}$$

for x as above and $y = w_0 + w_1 j_1 + w_2 j_2$, $w_0, w_1, w_2 \in A$. The associated cubic Jordan algebra will be denoted by $J(A, \mu) := J(X)$. The Jordan algebra A^+ identifies with a cubic subalgebra of $J(A, \mu)$ through the initial summand, and if A is central simple, then $J(A, \mu)$ is an Albert algebra, which is either split or division.

Now let (K, B, τ) be an associative algebra of degree 3 with unitary involution over F and suppose (u, μ) is a norm pair of (K, B, τ). If $K = F \times F$ is split, then we have canonical identifications $(B, \tau) = (A \times A^{\mathrm{op}}, \varepsilon)$ for some associative algebra A of degree 3 over F, where ε denotes the exchange involution, and $H(B, \tau) = A^+$ as cubic Jordan algebras, where the inclusion $H(B, \tau) \subseteq B$ corresponds to the diagonal embedding $A^+ \hookrightarrow A \times A^{\mathrm{op}}$. Moreover, $\mu = (\alpha, \beta)$, where $\alpha \in F$ is invertible, $\beta = \alpha^{-1} N_A(u)$, and there exists a canonical isomorphism $J := J(K, B, \tau, u, \mu) \cong J(A, \alpha) =: J'$ matching $H(B, \tau)$ canonically with A^+ as subalgebras of J, J', respectively. On the other hand, if K is a field, the preceding considerations apply to the base change from F to K and then yield an isomorphism $J(K, B, \tau, u, \mu)_K \cong J(B, \mu)$.

4 The weak and strong Skolem-Noether properties

As we have pointed out in 1.2, extending an isomorphism between cubic étale subalgebras of an Albert algebra J to an automorphism on all of J will in general not be possible. Working with elements of the structure group rather than automorphisms, our Theorem B above is supposed to serve as a substitute for this deficiency. Unfortunately, however, this substitute suffers from deficiencies of its own since the natural habitat of the structure group is the category of Jordan algebras *not* under homomorphisms *but, instead,* under homotopies.

Fixing a cubic Jordan algebra J over our base field F and a cubic étale F-algebra E throughout this section, we therefore feel justified in phrasing the following formal definition.

4.1 Weak and strong equivalence of isotopic embeddings

(a) Two isotopic embeddings $i, i' \colon E \to J$ in the sense of 2.8 are said to be *weakly equivalent* if there exist an element $w \in E$ of norm 1 and an element

$\psi \in \mathrm{Str}(J)$ such that the diagram

$$\begin{tikzcd} E \arrow[r, "R_w"'] \arrow[d, "i'"'] & E \arrow[d, "i"] \\ J \arrow[r, "\psi"'] & J \end{tikzcd} \tag{4.1.1}$$

commutes. They are said to be *strongly equivalent* if $\psi \in \mathrm{Str}(J)$ can furthermore be chosen so that the diagram commutes with $w = 1$ (i.e., $R_w = \mathrm{Id}_E$). Weak and strong equivalence clearly define equivalence relations on the set of isotopic embeddings from E to J.

(b) The pair (E, J) is said to satisfy the *weak (resp. strong) Skolem-Noether property for isotopic embeddings* if any two isotopic embeddings from E to J are weakly (resp. strongly) equivalent. The weak (resp. strong) Skolem-Noether property for isomorphic embeddings is defined similarly, by restricting the maps i, i' to be isomorphic embeddings instead of merely isotopic ones.

Remark 4.1.2. In 4.1 we have defined four different properties, depending on whether one considers the weak or strong Skolem-Noether property for isotopic or isomorphic embeddings. Clearly the combination weak/isomorphic is the weakest of these four properties and strong/isotopic is the strongest.

In the case where J is an Albert algebra, Theorem B is equivalent to saying that the pair (E, J) satisfies the weakest combination, the weak Skolem-Noether property for isomorphic embeddings. On the other hand, suppose $i, i' \colon E \to J$ are isomorphic embeddings and $\psi \in \mathrm{Str}(J)$ makes (4.1.1) commutative with $w = 1$. Then ψ fixes 1_J and hence is an automorphism of J. But such an automorphism will in general not exist [AJ, Th. 9], and if it doesn't the pair (E, J) will fail to satisfy the strong Skolem-Noether property for isomorphic embeddings. In view of this failure, we are led quite naturally to the following (as yet) open question:

Does the pair (E, J), with J absolutely simple (of degree 3), always satisfy the weak Skolem-Noether property for isotopic embeddings? (4.1.3)

This is equivalent to asking whether, given two cubic étale subalgebras $E_1 \subseteq J^{(p_1)}$, $E_2 \subseteq J^{(p_2)}$ for some $p_1, p_2 \in J^\times$, every isotopy $\eta \colon E_1 \to E_2$ allows a norm-one element $w \in E_1$ such that the isotopy $\eta \circ R_w \colon E_1 \to E_2$ extends to an element of the structure group of J. Regrettably, the methodological arsenal assembled in the present paper, consisting as it does of rather elementary manipulations involving the two Tits constructions, does not seem strong enough to provide an affirmative answer to this question.

But in the case where J is absolutely simple of dimension 9 — i.e., the Jordan algebra of symmetric elements in a central simple associative algebra of degree 3 with unitary involution over F [McCZ, 15.5] — we will show in Th. 5.2.7 below that the weak Skolem-Noether property for isotopic embeddings does hold. This result, in turn, will be instrumental in proving Theorem B in §7. Regarding the strong Skolem-Noether property for isomorphic embeddings, Theorem 1.1 in [GanS] gives a way to measure its failure.

5 Cubic Jordan algebras of dimension 9

Our goal in this section will be to answer Question 4.1.3 affirmatively in case J is a nine-dimensional absolutely simple cubic Jordan algebra over F. Before we will be able to do so, a few preparations are required.

5.1 Quadratic and cubic étale algebras

(a) If K and L are quadratic étale algebras over F, then so is

$$K * L := H(K \otimes L, \iota_K \otimes \iota_L),$$

where ι_K and ι_L denote the conjugations of K and L, respectively. The composition $(K, L) \mapsto K * L$ corresponds to the abelian group structure of $H^1(F, \mathbb{Z}/2\mathbb{Z})$, which classifies quadratic étale F-algebras [KMRT, (29.9)].
(b) Next suppose L and E are a quadratic and cubic étale F-algebras, respectively. Then $E \otimes L$ may canonically be viewed as a cubic étale L-algebra, whose norm, trace, adjoint will again be denoted by $N_E, T_E, \sharp$, respectively. On the other hand, $E \otimes L$ may also be viewed canonically as a quadratic étale E-algebra, whose norm, trace and conjugation will again be denoted by n_L, t_L, and ι_L, $x \mapsto \bar{x}$, respectively. We may and always will identify $E \subseteq E \otimes L$ through the first factor and then have $E = H(E \otimes L, \iota_L)$.

5.2 The étale Tits process

[PeT04a, 1.3] Let L, resp. E, be a quadratic, resp cubic, étale algebra over F and as in 3.3 write $\Delta(E)$ for the discriminant of E, which is a quadratic étale F-algebra. With the conventions of 5.1 (b), the triple $(K, B, \tau) := (L, E \otimes L, \iota_L)$ is an associative algebra of degree 3 with unitary involution over F in the sense of 3.7 such that $H(B, \tau) = E$. Hence, if (u, b) is a norm pair of $(L, E \otimes L, \iota_L)$, the second Tits construction 3.8 leads to a cubic Jordan algebra

$$J(E, L, u, b) := J(K, B, \tau, u, b) = J(L, E \otimes L, \iota_L, u, b)$$

that belongs to the cubic norm structure $(V, c, \sharp, N)$ where $V = E \oplus (E \otimes L)j$ as a vector space over F and $c, \sharp, N$ are defined by (3.8.1)–(3.8.3) in all scalar extensions. The cubic Jordan algebra $J(E, L, u, b)$ is said to arise from E, L, u, b by means of the *étale Tits process*. There exists a central simple associative algebra (B, τ) of degree 3 with unitary involution over F uniquely determined by the condition that $J(E, L, u, b) \cong H(B, \tau)$, and by [PeR 84b, Th. 1], the centre of B is isomorphic to $\Delta(E) * L$ (cf. 5.1 (a)) as a quadratic étale F-algebra.
For convenience, we now recall three results from [PeT04a] that will play a crucial role in providing an affirmative answer to Question 4.1.3 under the conditions spelled out at the beginning of this section.

Theorem 5.2.1. ([PeT04a, 1.6]) *Let E be a cubic étale F-algebra, (B, τ) a central simple associative algebra of degree 3 with unitary involution over F and suppose i is an isomorphic embedding from E to $H(B, \tau)$. Writing K for the centre of B and setting $L := K * \Delta(E)$, there is a norm pair (u, b) of $(L, E \otimes L, \iota_L)$ such that i extends to an isomorphism from the étale Tits process algebra $J(E, L, u, b)$ onto $H(B, \tau)$.* □

Theorem 5.2.2. ([PeT04a, 3.2]) *Let E, E' and L, L' be cubic and quadratic étale algebras, respectively, over F and suppose we are given norm pairs (u, b) of $(L, E \otimes L, \iota_L)$ and (u', b') of $(L', E' \otimes L', \iota_{L'})$. We write*

$$J := J(E, L, u, b) = E \oplus (E \otimes L)j, \quad J' := J(E', L', u', b') = E' \oplus (E' \otimes L')j'$$

as in 5.2 for the corresponding étale Tits process algebras and let $\varphi\colon E' \xrightarrow{\sim} E$ be an isomorphism. Then, for an arbitrary map $\Phi\colon J' \to J$, the following conditions are equivalent.

(i) Φ *is an isomorphism extending* φ.

(ii) There exist an isomorphism $\psi\colon L' \xrightarrow{\sim} L$ and an invertible element $y \in E \otimes L$ such that $\varphi(u') = n_L(y)u$, $\psi(b') = N_E(y)b$ and

$$\Phi(v_0' + v'j') = \varphi(v_0') + \big(y(\varphi \otimes \psi)(v')\big)j \tag{5.2.3}$$

for all $v_0' \in E'$, $v' \in E' \otimes L'$. □

Proposition 5.2.4. ([PeT04a, 4.3]) *Let E be a cubic étale F-algebra and $\alpha, \alpha' \in F^\times$. Then the following conditions are equivalent.*

i. The first Tits constructions $J(E, \alpha)$ and $J(E, \alpha')$ (cf. 3.9) are isomorphic.

ii. $J(E, \alpha)$ and $J(E, \alpha')$ are isotopic.

iii. $\alpha \equiv \alpha'^{\varepsilon} \bmod N_E(E^\times)$ for some $\varepsilon = \pm 1$.

iv. The identity of E can be extended to an isomorphism from $J(E, \alpha)$ onto $J(E, \alpha')$. □

Our next aim will be to derive a version of Th. 5.2.1 that works with isotopic rather than isomorphic embeddings and brings in a normalization condition already known from [KMRT, (39.2)].

Proposition 5.2.5. *Let (B, τ) be a central simple associative algebra of degree 3 with unitary involution over F and write K for the centre of B. Suppose further that E is a cubic étale F-algebra and put $L := K * \Delta(E)$. Given any isotopic embedding $i\colon E \to J := H(B, \tau)$, there exist elements $u \in E$, $b \in L$ such that $N_E(u) = n_L(b) = 1$ and i can be extended to an isotopy from $J(E, L, u, b)$ onto J.*

Proof. By 2.8, some invertible element $p \in J$ makes $i\colon E \to J^{(p)}$ an isomorphic embedding. On the other hand, invoking 2.7 and writing $\tau^{(p)}$ for the p-twist of τ, it follows that

$$R_p\colon J^{(p)} \xrightarrow{\sim} H(B, \tau^{(p)})$$

is an isomorphism of cubic Jordan algebras, forcing $i_1 := R_p \circ i\colon E \to H(B, \tau^{(p)})$ to be an isomorphic embedding. Hence Th. 5.2.1 yields a norm pair (u_1, μ_1) of $(L, E \otimes L, \iota_L)$ such that, adapting the notation of 3.8 to the present set-up in an obvious manner, i_1 extends to an isomorphism

$$\eta_1'\colon J(E, L, u_1, b_1) = E \oplus (E \otimes L)j_1 \xrightarrow{\sim} H(B, \tau^{(p)}).$$

Thus $\eta_1 := R_{p^{-1}} \circ \eta_1'\colon J(E, L.u_1, b_1) \xrightarrow{\sim} J^{(p)}$ is an isomorphism, which may therefore be viewed as an isotopy from $J(E, L, u_1, b_1)$ onto J extending i. Now put $u := N_E(u_1)^{-1}u_1^3$, $b := \bar{b}_1 b_1^{-1}$ and $y := u_1 \otimes b_1^{-1} \in (E \otimes L)^\times$ to conclude $N_E(u) = n_L(b) = 1$ as well as $n_L(y)u_1 = u$, $N_E(y)b_1 = b$. Applying Th. 5.2.2 to $\varphi := \mathbf{1}_E$, $\psi := \mathbf{1}_L$, we therefore obtain an isomorphism

$$\Phi\colon J(E, L, u, b) \xrightarrow{\sim} J(E, L, u_1, b_1), \quad v_0 + vj_1 \longmapsto v_0 + (yv)j$$

of cubic Jordan algebras, and $\eta := \eta_1 \circ \Phi\colon J(E, L, u, b) \to J$ is an isotopy of the desired kind. □

Lemma 5.2.6. *Let L, resp. E be a quadratic, resp. cubic étale algebra over F and suppose we are given elements $u \in E$, $b \in L$ satisfying $N_E(u) = n_L(b) = 1$. Then $w := u^{-1} \in E$ has norm 1 and $R_w\colon E \to E$ extends to an isomorphism*

$$\hat{R}_w\colon J(E, L, 1_E, b) \xrightarrow{\sim} J(E, L, u, b)^{(u)}, \quad v + xj \longmapsto (vw) + xj$$

of cubic Jordan algebras.

Proof. This follows immediately from Lemma 3.8.5 for $(K, B, \tau) := (L, E \otimes L, \iota_L)$, $\mu := b$ and $p := u$. □

We are now ready for the main result of this section.

Theorem 5.2.7. *Let (B, τ) be a central simple associative algebra of degree 3 with unitary involution over F and E a cubic étale F-algebra. Then the pair (E, J) with $J := H(B, \tau)$ satisfies the weak Skolem-Noether property for isotopic embeddings in the sense of 4.1 (b).*

Proof. Given two isotopic embeddings $i, i'\colon E \to J$, we must show that they are weakly equivalent. In order to do so, we write K for the centre of B as a quadratic étale F-algebra and put $L := K * \Delta(E)$. Then Prop. 5.2.5 yields elements $u, u' \in E$, $b, b' \in L$ satisfying

$$N_E(u) = N_E(u') = n_L(b) = n_L(b') = 1 \tag{5.2.8}$$

such that the isotopic embeddings i, i' can be extended to isotopies

$$\begin{aligned} \eta\colon\ J(E,L,u,b) &= E \oplus (E \otimes L)j \longrightarrow J, \\ \eta'\colon\ J(E,L,u',b') &= E \oplus (E \otimes L)j' \longrightarrow J, \end{aligned} \tag{5.2.9}$$

respectively. We now distinguish the following two cases.
Case 1: $L \cong F \times F$ is split. As we have noted in 3.9, there exist elements $\alpha, \alpha' \in F^\times$ and isomorphisms

$$\Phi\colon\ J(E,L,u,b) \xrightarrow{\sim} J(E,\alpha), \quad \Phi'\colon\ J(E,L,u',b') \xrightarrow{\sim} J(E,\alpha')$$

extending the identity of E. Thus (5.2.9) implies that $\Phi \circ \eta^{-1} \circ \eta' \circ \Phi'^{-1}\colon\ J(E,\alpha') \to J(E,\alpha)$ is an isotopy, and applying Prop. 5.2.4, we find an isomorphism $\theta\colon\ J(E,\alpha') \xrightarrow{\sim} J(E,\alpha)$ extending the identity of E. But then $\varphi := \eta \circ \Phi^{-1} \circ \theta \circ \Phi' \circ \eta'^{-1}\colon\ J \longrightarrow J$ is an isotopy, hence belongs to the structure group of J, and satisfies

$$\varphi \circ i' = \eta \circ \Phi^{-1} \circ \theta \circ \Phi' \circ \eta'^{-1} \circ \eta'|_E = \eta|_E = i.$$

Thus i and i' are even strongly equivalent.
Case 2: L is a field. Since $J(E,L,u,b)$ and $J(E,L,u',b')$ are isotopic (via $\eta'^{-1} \circ \eta$), so are their scalar extensions from F to L. From this and 3.9 we therefore conclude that $J(E \otimes L, b)$ and $J(E \otimes L, b')$ are isotopic over L. Hence, by Prop. 5.2.4,

$$b = b'^{\varepsilon} N_E(z) \tag{5.2.10}$$

for some $\varepsilon = \pm 1$ and some $z \in (E \otimes L)^\times$. Now put $\varphi := \mathbf{1}_E$, $\psi := \iota_L$ and $y := u' \otimes 1_L \in (E \otimes L)^\times$. Making use of (5.2.8) we deduce $n_L(y)u'^{-1} = u'$, $N_E(y)b'^{-1} = \bar{b}'$. Hence Th. 5.2.2 shows that the identity of E can be extended to an isomorphism

$$\theta\colon\ J(E,L,u',b') \xrightarrow{\sim} J(E,L,u'^{-1},b'^{-1}),$$

and we still have $N_E(u'^{-1}) = n_L(b'^{-1}) = 1$. Thus, replacing η' by $\eta' \circ \theta^{-1}$ if necessary, we may assume $\varepsilon = 1$ in (5.2.10), i.e.,

$$b = b' N_E(z). \tag{5.2.11}$$

Next put $\varphi := \mathbf{1}_E$, $\psi := \mathbf{1}_L$ and $y := z \in (E \otimes L)^\times$, $u_1 := n_L(y)u'$, $b_1 := N_E(y)b' = b$ (by (5.2.11)). Taking L-norms in (5.2.11) and observing (5.2.8), we conclude $N_E(y)\overline{N_E(y)} = n_L\big(N_E(z)\big) = 1$, and since $u_1 = y\bar{y}u'$, this implies $N_E(u_1) = 1$. Hence Th. 5.2.2 yields an isomorphism

$$\theta\colon\ J(E,L,u_1,b_1) \xrightarrow{\sim} J(E,L,u',b')$$

extending the identity of E, and replacing η' by $\eta' \circ \theta$ if necessary, we may and from now on will assume

$$b = b'. \tag{5.2.12}$$

Setting $w := u^{-1}$ and consulting Lemma 5.2.6, we have $N_E(w) = 1$ and obtain a commutative diagram

$$\begin{tikzcd} E \arrow[r, "R_w"] \arrow[d, hook] & E \arrow[r, "i"] \arrow[d, hook] & J \\ J(E,L,1_E,b) \arrow[r, "\hat{R}_w"'] & J(E,L,u,b), \arrow[ur, "\eta"'] & \end{tikzcd}$$

where $\eta \circ \hat{R}_w\colon\ J(E,L,1_E,b) \to J$ is an isotopy and the isotopic embeddings $i, i \circ R_w$ from E to J are easily seen to be weakly equivalent. Hence, replacing i by $i \circ R_w$ and η by $\eta \circ \hat{R}_w$ if necessary, we may assume $u = 1_E$. But then, by symmetry, we may assume $u' = 1_E$ as well, forcing

$$\eta, \eta'\colon\ J(E,L,1_E,b) \longrightarrow J$$

to be isotopies extending i, i', respectively. Thus $\psi := \eta \circ \eta'^{-1} \in \mathrm{Str}(J)$ satisfies $\psi \circ i' = \eta \circ \eta'^{-1} \circ \eta'|_E = \eta|_E = i$, so i and i' are strongly, hence weakly, equivalent. □

6 Norm classes and strong equivalence

6.1

Let (B, τ) be a central simple associative algebra of degree 3 with unitary involution over F and E a cubic étale F-algebra. Then the centre, K, of B and the discriminant, $\Delta(E)$, of E are quadratic étale F-algebras, as is $L := K * \Delta(E)$ (cf. 5.1 (a)). To any pair (i, i') of isotopic embeddings from E to $J := H(B, \tau)$ we will attach an invariant, belonging to $E^\times / n_L((E \otimes L)^\times)$ and called the norm class of (i, i'), and we will show that i and i' are strongly equivalent if and only if their norm class is trivial. In order to achieve these objectives, a number of preparations will be needed.
We begin with an extension of Th. 5.2.2 from isomorphisms to isotopies.

Proposition 6.1.1. *Let E, E' and L, L' be cubic and quadratic étale algebras, respectively, over F and suppose we are given norm pairs (u, b) of $(L, E \otimes L, \iota_L)$ and (u', b') of $(L', E' \otimes L', \iota_{L'})$. We write*

$$J := J(E,L,u,b) = E \oplus (E \otimes L)j, \quad J' := J(E',L',u',b') = E' \oplus (E' \otimes L')j'$$

as in 5.2 for the corresponding étale Tits process algebras and let $\varphi\colon\ E' \xrightarrow{\sim} E$ be an isotopy. Then, letting $\Phi\colon\ J' \to J$ be an arbitrary map and setting $p := \varphi(1_{E'})^{-1} \in E^\times$, the following conditions are equivalent.

i. Φ is an isotopy extending φ.

ii. There exist an isomorphism $\psi\colon\ L' \xrightarrow{\sim} L$ and an invertible element $y \in E \otimes L$ such that $\varphi(u') = n_L(y)p^\sharp p^{-3}u$, $\psi(b') = N_E(y)b$ and

$$\Phi(v_0' + v'j') = \varphi(v_0') + \big(y(\varphi \otimes \psi)(v')\big)j \tag{6.1.2}$$

for all $v_0' \in E'$, $v' \in E' \otimes L'$.

Proof. $\varphi_1 := R_p \circ \varphi \colon E' \to E$ is an isotopy preserving units, hence is an isomorphism. By 5.2 we have

$$J := J(E, L, u, b) = J(L, E \otimes L, \iota_L, u, b),$$

and in obvious notation, setting $u^{(p)} := p^\sharp u$, $b^{(p)} := N_E(p)b$, Lemma 3.8.5 yields an isomorphism

$$\begin{aligned} \hat{R}_p \colon J^{(p)} \xrightarrow{\sim} J_1 &:= J(L, E \otimes L, \iota_L, u^{(p)}, b^{(p)}) = J(E, L, u^{(p)}, b^{(p)}), \\ v_0 + vj &\longmapsto (v_0 p) + vj_1 \end{aligned}$$

Thus $\hat{R}_p \colon J \to J_1$ is an isotopy and $\Phi_1 := \hat{R}_p \circ \Phi$ is a map from J' to J_1. Since φ_1 preserves units, this leads to the following chain of equivalent conditions.

$$\begin{aligned} \Phi \text{ is an isotopy extending } \varphi &\Longleftrightarrow \Phi_1 \text{ is an isotopy extending } \varphi_1 \\ &\Longleftrightarrow \Phi_1 \text{ is an isotopy extending } \varphi_1 \\ &\qquad\;\; \text{and preserving units} \\ &\Longleftrightarrow \Phi_1 \text{ is an isomorphism extending } \varphi_1. \end{aligned}$$

By Th. 5.2.2, therefore, (i) holds if and only if there exist an element $y_1 \in (E \otimes L)^\times$ and an isomorphism $\psi \colon L' \to L$ such that $\varphi_1(u') = n_L(y_1)u^{(p)}$, $\psi(b') = N_E(y_1)b^{(p)}$ and

$$\Phi_1(v_0' + v'j') = \varphi_1(v_0') + \big(y_1(\varphi_1 \otimes \psi)(v')\big)j_1$$

for all $v_0' \in E'$, $v' \in E' \otimes L'$. Setting $y := y_1 p$, and observing $(\varphi_1 \otimes \psi)(v') = (\varphi \otimes \psi)(v')p$ for all $v' \in E' \otimes L'$, it is now straightforward to check that the preceding equations, in the given order, are equivalent to the ones in condition (ii) of the theorem. □

With the notational conventions of 5.1 (b), we next recall the following result.

Lemma 6.1.3. ([PeT04a, Lemma 4.5]) *Let L (resp. E) be a quadratic (resp. a cubic) étale F-algebra. Given $y \in E \otimes L$ such that $c := N_E(y)$ satisfies $n_L(c) = 1$, there exists an element $y' \in E \otimes L$ satisfying $N_E(y') = c$ and $n_L(y') = 1$.* □

6.2 Notation

For the remainder of this section we fix a central simple associative algebra (B, τ) of degree 3 with unitary involution over F and a cubic étale F-algebra E. We write K for the centre of B, put $J := H(B, \tau)$ and $L := K * \Delta(E)$ in the sense of 5.1.

Theorem 6.2.1. *Let $i \colon E \to J$ be an isotopic embedding and suppose $w \in E$ has norm 1. Then the isotopic embeddings i and $i \circ R_w$ from E to J are strongly equivalent if and only if $w \in n_L((E \otimes L)^\times)$.*

Proof. By Prop. 5.2.5, we find a norm pair (u, b) of $(L, E \otimes L, \iota_L)$ such that i extends to an isotopy $\eta\colon J_1 := J(E, L, u, b) \to J$. On the other hand, i and $i \circ R_w$ are strongly equivalent by definition (cf. 4.1) if and only if there exists an element $\Psi \in \mathrm{Str}(J)$ making the central square in the diagram

$$\begin{tikzcd} J_1 \arrow[dr, "\eta"'] & E \arrow[l, hook'] \arrow[r, "R_w"'] \arrow[d, "i"'] & E \arrow[r, hook] \arrow[d, "i"] & J_1 \arrow[dl, "\eta"] \\ & J \arrow[r, dotted, "\Psi"'] & J & \end{tikzcd} \qquad (6.2.2)$$

commutative, equivalently, the isotopy $\varphi := R_w\colon E \to E$ can be extended to an element of the structure group of J_1. By Prop. 6.1.1 (with $p = w^{-1}$), this in turn happens if and only if some invertible element $y \in E \otimes L$ has $uw = n_L(y)(w^{-1})^\sharp w^3 u = n_L(y)w^4 u$, i.e., $w = n_L(w^2 y)$, and either$N_E(y) = 1$ or $N_E(y) = \bar{b}b^{-1}$. Replacing y by $w^2 y$, we conclude that i and $i \circ R_w$ are strongly equivalent if and only

$$\text{some } y \in E \otimes L \text{ satisfies (i) } n_L(y) = w \text{ and (ii) } N_E(y) \in \{1, \bar{b}b^{-1}\}. \qquad (6.2.3)$$

In particular, for i and $i \circ R_w$ to be strongly equivalent it is necessary that $w \in n_L((E \otimes L)^\times)$. Conversely, let this be so. Then some $y \in E \otimes L$ satisfies condition (i) of (6.2.3), so we have $w = n_L(y)$ and $n_L(N_E(y)) = N_E(n_L(y)) = N_E(w) = 1$. Hence Lemma 6.1.3 yields an element $y' \in E \otimes L$ such that $N_E(y') = N_E(y)$ and $n_L(y') = 1$. Setting $z := yy'^{-1} \in E \otimes L$, we conclude $n_L(z) = n_L(y) = w$ and $N_E(z) = N_E(y)N_E(y')^{-1} = 1$, hence that (6.2.3) holds for z in place of y. Thus i and $i \circ R_w$ are strongly equivalent. □

6.3 Norm classes

Let $i, i'\colon E \to J$ be isotopic embeddings. By Th. 5.2.7, there exist a norm-one element $w \in E$ as well as an element $\psi \in \mathrm{Str}(J)$ such that the left-hand square of the diagram

$$\begin{tikzcd} E \arrow[r, "R_w"'] \arrow[d, "i'"'] & E \arrow[d, "i"] & E \arrow[l, "R_{w'}"] \arrow[d, "i'"] \\ J \arrow[r, "\psi"'] & J & J \arrow[l, "\psi'"] \end{tikzcd}$$

commutes. Given another norm-one element $w' \in E$ and another element $\psi' \in \mathrm{Str}(J)$ such that the right-hand square of the above diagram commutes as well, then the isotopic embeddings i' and $i' \circ R_{ww'^{-1}}$ from E to J are strongly equivalent (via $\psi'^{-1} \circ \psi$), and Th. 6.2.1 implies $w \equiv w' \bmod n_L((E \otimes L)^\times)$. Thus the class of $w \bmod n_L((E \otimes L)^\times)$ does not depend on the choice of w and ψ. We write this class as $[i, i']$ and call it the *norm class* of (i, i'); it is clearly symmetric in i, i'. We say i, i' *have trivial norm class* if

$$[i, i'] = 1 \text{ in } E^\times / n_L((E \otimes L)^\times).$$

For three isotopic embeddings $i, i', i''\colon E \to J$, it is also trivially checked that $[i, i''] = [i, i'][i', i'']$.

Corollary 6.3.1. *Two isotopic embeddings $i, i'\colon E \to J$ are strongly equivalent if and only if $[i, i']$ is trivial.*

Proof. Let $i, i'\colon E \to J$ be isotopic embeddings. By Th. 5.2.7, they are weakly equivalent, so some norm-one element $w \in E$ makes i' and $i \circ R_w$ strongly equivalent. Thus i and i' are strongly equivalent if and only if i and $i \circ R_w$ are strongly equivalent, which by Th. 6.2.1 amounts to the same as $w \in n_L((E \otimes L)^\times)$, i.e., to i and i' having trivial norm class. ☐

Remark 6.3.2. When confined to isomorphic rather than isotopic embeddings, Cor. 6.3.1 reduces to [PeT04a, Th. 4.2].

6.4 The connection with Jordan pairs.

We are grateful to a referee for having suggested to phrase some of the preceding results in the language of Jordan pairs. Her or his arguments may be sketched as follows.
Referring to Loos [Loos] for the necessary background material, we consider cubic étale F-algebras E, E', a central simple associative algebra (B, τ) of degree 3 with unitary involution over F, put $J := H(B, \tau)$ and write $\mathcal{E} := (E^+, E^+)$, $\mathcal{E}' := (E'^+, E'^+)$, $\mathcal{J} := (J, J)$ for the Jordan pairs corresponding to the Jordan algebra E^+, E'^+, J, respectively. Then we make the following observations.

1^0. *Isotopic embeddings from E to J are basically the same as embeddings* (i.e., injective homomorphisms) *of Jordan pairs from $\mathcal{E}$ to $\mathcal{J}$.* Indeed, let $\varphi\colon E^+ \to J$ be an isotopic embedding, so some $p \in J^\times$ makes $\varphi\colon E^+ \to J^{(p)}$ an injective homomorphism. Arguing as in the proof of [Loos, 1.8], it then follows that $(\varphi, U_p\varphi)\colon \mathcal{E} \to \mathcal{J}$ is an embedding of Jordan pairs. Conversely, let $h = (h_+, h_-)\colon \mathcal{E} \to \mathcal{J}$ be an embedding of Jordan pairs.. Then $e := (e_+, e_-) := (h_+(1_J), h_-(1_J))$ is an idempotent of $\mathcal{J}$ such that $h(\mathcal{E}) \subseteq \mathcal{J}_2(e)$. We will be through once we have shown $\mathcal{J}_2(e) = \mathcal{J}$ since this implies, again following the proof of loc.cit., that $p := e_-$ is invertible in J and h_+ is an injective homomorphism from from E^+ to $J^{(p)}$, hence an isotopic embedding from E to J. In order to show $\mathcal{J}_2(e) = \mathcal{J}$, we may extend scalars if necessary to the algebraic closure of F to obtain a frame $X = (e_1, e_2, e_3)$ in $\mathcal{J}_2(e)$ satisfying $e = \sum e_i$. But since $\mathcal{J}$ has degree three, X is also a frame of $\mathcal{J}$, forcing $e \in \mathcal{J}$ to be a maximal idempotent. Now [Pe 78, Cor. 4] implies the assertion.

2^0. Let $w \in E^\times$. Then it is straightforward to check that $R_w\colon E \to E$ *belongs the structure group of E^+, and the automorphism of $\mathcal{E}$ it corresponds to via [Loos, 1.8] is $\mathcal{R}_w := (R_w, R_{w^{-1}})$.*

3^0. We claim that *the isomorphisms from $\mathcal{E}$ to $\mathcal{E}'$ are precisely the pairs of maps $(\varphi \circ R_w, \varphi \circ R_{w^{-1}}) = (\varphi, \varphi) \circ \mathcal{R}_w$, where $\varphi\colon E \to E'$ is an isomorphism of cubic étale F-algebras and $w \in E^\times$.* By 2^0, maps of this form are clearly isomorphisms from $\mathcal{E}$ to $\mathcal{E}'$. Conversely, let $h = (h_+, h_-)\colon \mathcal{E} \to \mathcal{E}'$ be an isomorphism. Then $v := h_+^{-1}(1_{E'}) \in E^\times$ and $\varphi := h_+ \circ R_{w^{-1}}\colon E^+ \to E'^+$ with $w := v^{-1}$ is an isotopy preserving units, hence an isomorphism of cubic étale

F-algebras (Lemma 2.5.2) such that $h_+ = \varphi \circ R_w$. Now [Loos, 1.8] yields also $h_- = \varphi \circ R_{w^{-1}}$.

4^0. We denote by $\mathrm{EAut}(\mathcal{J})$ the *elementary automorphism group* of $\mathcal{J}$, meaning the group of automorphisms (h_+, h_-) of $\mathcal{J}$ such that both h_+ and h_- leave the norm of J invariant. Under the correspondence of [Loos, 1.8], $\mathrm{EAut}(\mathcal{J})$ identifies canonically with the *elementary structure group* of J, denoted by $\mathrm{EStr}(J)$ and defined to be the group of all $g \in \mathrm{Str}(J)$ that leave the norm of J invariant.

5^0. We are now in a position to phrase, e.g., the weak Skolem-Noether property for isotopic embeddings (cf. 4.1) in the language of Jordan pairs as follows: *given any two isomorphic embeddings $j, j' \colon \mathcal{E} \to \mathcal{J}$, there exist an invertible element $w \in E$ and an elementary automorphism Ψ of $\mathcal{J}$ such that the diagram*

$$\begin{tikzcd} \mathcal{E} \arrow[r, "\mathcal{R}_w"'] \arrow[d, "j'"'] & \mathcal{E} \arrow[d, "j"] \\ \mathcal{J} \arrow[r, "\Psi"'] & \mathcal{J} \end{tikzcd}$$

commutes. Note that we have replaced the normalization condition $N_E(w) = 1$ of 4.1 by the equivalent one of Ψ being an elementary (rather than arbitrary) automorphism of $\mathcal{J}$. Combining Th. 5.2.7 with 3^0 we also see that *two maximal tori of $\mathcal{J}$ are conjugate under* $\mathrm{EAut}(\mathcal{J})$ *if and only if they are isomorphic.*

7 Albert algebras: proof of Theorem B

7.1

It would be interesting to know whether Th. 5.2.7, the notion of norm class as defined in 6.3, or Cor. 6.3.1 can be extended from absolutely simple Jordan algebras of degree 3 and dimension 9 to Albert algebras. Unfortunately, we have neither been able either to answer this question in the affirmative nor to exhibit a counter example. Therefore, we will have to be more modest by settling with Theorem B, i.e., with the weak Skolem-Noether property for isomorphic rather than arbitrary isotopic embeddings. Given a cubic étale algebra E and an Albert algebra J over F, the idea of the proof is to factor two isomorphic embeddings from E to J through the same absolutely simple nine-dimensional subalgebra of J, which by structure theory will have the form $H(B, \tau)$ for some central simple associative algebra (B, τ) of degree 3 with unitary involution over F, allowing us to apply Th. 5.2.7 and reach the desired conclusion. The fact that we have not succeeded in extending the preceding factorization property from isomorphic to isotopic embeddings from E to J is the main reason for the deficiencies alluded to at the beginning of this section. Throughout, we fix an arbitrary Albert algebra J and a cubic étale algebra E over F. In order to carry out the procedure we have just described, a few preparations will be needed.

Lemma 7.1.1. *Assume F is algebraically closed and denote by $E_1 := \mathrm{Diag}_3(F) \subseteq \mathrm{Mat}_3(F)^+$ the cubic étale subalgebra of diagonal matrices. Then there exists a cubic étale subalgebra $E_2 \subseteq \mathrm{Mat}_3(F)^+$ such that $\mathrm{Mat}_3(F)^+$ is generated by E_1 and E_2 as a cubic Jordan algebra over F.*

Proof. We realize $\mathrm{Mat}_3(F)^+$ as a first Tits construction

$$J_1 := \mathrm{Mat}_3(F)^+ = J(E_1, 1),$$

with adjoint $\sharp$, norm N, trace T, and identify the diagonal matrices on the left with E_1 viewed canonically as a cubic subalgebra of $J(E_1, 1)$. Since F is infinite, we find an element $u_0 \in E_1$ satisfying $E_1 = F[u_0]$. Letting $\alpha \in F^\times$, we put

$$y := u_0 + \alpha j_1 \in J_1.$$

Since u_0 and j_1 generate J_1 as a cubic Jordan algebra, so do u_0 and y, hence E_1 and $E_2 := F[y]$. It remains to show that, for a suitable choice of α, the F-algebra E_2 is cubic étale. We first deduce from (3.9.1) and (3.9.3) that

$$\begin{aligned} y^\sharp &= u_0^\sharp + (-\alpha u_0)j_1 + \alpha^2 j_2, \\ T(y) &= T_{E_1}(u_0), \\ T(y^\sharp) &= T_{E_1}(u_0^\sharp), \\ N(y) &= N_{E_1}(u_0) + \alpha^3. \end{aligned}$$

Thus y has the generic minimum (= characteristic) polynomial

$$\mathbf{t}^3 - T_{E_1}(u_0)\mathbf{t}^2 + T_{E_1}(u_0^\sharp)\mathbf{t} - \big(N_{E_1}(u_0) + \alpha^3\big) \in F[\mathbf{t}],$$

whose discriminant by [Lang, IV, Exc. 12(b)] is

$$\begin{aligned} \Delta_y := {} & T_{E_1}(u_0)^2 T_{E_1}(u_0^\sharp)^2 - 4T_{E_1}(u_0^\sharp)^3 - 4T_{E_1}(u_0)^3(N_{E_1}(u_0) + \alpha^3) \\ & - 27(N_{E_1}(u_0) + \alpha^3)^2 + 18T_{E_1}(u_0)T_{E_1}(u_0^\sharp)(N_{E_1}(u_0) + \alpha^3) \\ = {} & \Delta_{u_0} - \big(4T_{E_1}(u_0)^3 + 54N_{E_1}(u_0) - 18T_{E_1}(u_0)T_{E_1}(u_0^\sharp)\big)\alpha^3 - 27\alpha^6, \end{aligned}$$

where $\Delta_{u_0} \neq 0$ is the discriminant of the minimum polynomial of u_0. Regardless of the characteristic, we can therefore choose $\alpha \in F^\times$ in such a way that $\Delta_y \neq 0$, in which case E_2 is a cubic étale F-algebra. ∎

7.2 Digression: pointed quadratic forms

By a *pointed quadratic form* over F we mean a triple (V, q, c) consisting of an F-vector space V, a quadratic form $q\colon V \to F$, with bilinearization $q(x, y) = q(x + y) - q(x) - q(y)$, and an element $c \in V$ that is a *base point* for q in the sense that $q(c) = 1$. Then V together with the U-operator

$$U_x y := q(x, \bar{y})x - q(x)\bar{y} \qquad (x, y \in V), \tag{7.2.1}$$

where $\bar{y} := q(c,y)c - y$, and the unit element $1_J := c$ is a Jordan algebra over F, denoted by $J := J(V,q,c)$ and called the Jordan algebra of the pointed quadratic form (V,q,c). It follows immediately from (7.2.1) that the subalgebra of J generated by a family of elements $x_i \in J$, $i \in I$, is $Fc + \sum_{i\in I} Fx_i$.

Lemma 7.2.2. *Assume F is infinite and let $i, i' \colon E \to J$ be isomorphic embeddings. Then there exist isomorphic embeddings $i_1, i'_1 \colon E \to J$ such that i (resp., i') is strongly equivalent to i_1 (resp., i'_1) and the subalgebra of J generated by $i_1(E) \cup i'_1(E)$ is absolutely simple of degree 3 and dimension 9.*

Proof. We proceed in two steps. Assume first that F is algebraically closed. Then $E = F \times F \times F$ and $J = \mathrm{Her}_3(C)$ are both split, C being the octonion algebra of Zorn vector matrices over F. Note that $\mathrm{Mat}_3(F)^+ \cong \mathrm{Her}_3(F \times F)$ may be viewed canonically as a subalgebra of J. By splitness of E, there are frames (i.e., complete orthogonal systems of absolutely primitive idempotents) $(e_p)_{1\le p\le 3}$, $(e'_p)_{1\le p\le 3}$ in J such that $i(E) = \sum Fe_p$, $i'(E) = \sum Fe'_p$. But frames in the split Albert algebra are conjugate under the automorphism group. Hence we find automorphisms φ, ψ of J satisfying $\varphi(e_p) = \psi(e'_p) = e_{pp}$ for $1 \le p \le 3$. Applying Lemma 7.1.1, we find a cubic étale subalgebra $E_2 \subseteq \mathrm{Mat}_3(F)^+ \subseteq J$ that together with $E_1 := \mathrm{Diag}_3(F) = (\varphi \circ i)(E)$ generates $\mathrm{Mat}_3(F)^+$ as a cubic Jordan algebra over F. Again, the cubic étale E_2 is split, so we find a frame $(c_p)_{1\le p\le 3}$ in J satisfying $E_2 = \sum Fc_p$. This in turn leads to an automorphism ψ' of J sending e_{pp} to c_p for $1 \le p \le 3$. Then $i_1 := \varphi \circ i$ and $i'_1 := \psi' \circ \psi \circ i'$ are strongly equivalent to i, i', respectively, and satisfy $i_1(E) = E_1$, $i'_1(E) = E_2$, hence have the desired property.

Now let F be an arbitrary infinite field and write $\bar{F}$ for its algebraic closure. We have $E = F[u]$ for some $u \in E$ and put $x := i(u), x' := i'(u) \in J$. We write k-**alg** for the category of commutative associative k-algebras with 1, put $G := \mathrm{Aut}(J) \times \mathrm{Aut}(J)$ as a group scheme over F and, given $R \in k$-**alg**, $(\varphi, \varphi') \in G(R)$, write $x_m := x_m(\varphi, \varphi')$, $1 \le m \le 9$, in this order for the elements

$$x_1 := 1_{J_R}, \quad x_2 := \varphi(x_R), \quad x_3 := \varphi(x_R^\sharp),$$
$$x_4 := \varphi'(x_R), \quad x_5 := \varphi'(x_R^\sharp), \quad x_6 := \varphi(x_R) \times \varphi'(x_R),$$
$$x_7 := \varphi(x_R^\sharp) \times \varphi'(x_R), \quad x_8 := \varphi(x_R) \times \varphi'(x_R^\sharp), \quad x_9 := \varphi(x_R^\sharp) \times \varphi'(x^\sharp).$$

By a result of Brühne (cf. [Pe 15, Prop. 6.6]), the subalgebra of J_R generated by $(\varphi \circ i_R)(E_R)$ and $(\varphi' \circ i'_R)(E_R)$ is spanned as an R-module by the elements $x_1, \ldots, x_9$. Now consider the open subscheme $X \subseteq G$ defined by the condition that $X(R)$, $R \in k$-**alg**, consist of all elements $(\varphi, \varphi') \in G(R)$ satisfying

$$\det\Big(T_J\big(x_m(\varphi,\varphi'), x_n(\varphi,\varphi')\big)\Big)_{1\le m,n\le 9} \in R^\times.$$

By what we have just seen, this is equivalent to saying that the subalgebra of J_R generated by $(\varphi \circ i_R)(E_R)$ and $(\varphi' \circ i'_R)(E_R)$ is a free R-module of rank 9 and has a non-singular trace form. By the preceding paragraph, $X(\bar{F}) \subseteq G(\bar{F})$ is a non-empty (Zariski-) open, hence dense, subset. On the other hand, by [Sp, 13.3.9(iii)], $G(F)$ is dense in $G(\bar{F})$. Hence so is $X(F) = X(\bar{F}) \cap G(F)$. In particular, we can find elements $\varphi, \varphi' \in \mathrm{Aut}(J)(F)$ such that the subalgebra J' of J generated by $(\varphi \circ i)(E)$ and $(\varphi' \circ i')(E)$ is non-singular of dimension 9. This property is preserved under base field extensions, as is the property of being generated by two elements. Hence, if J' were not absolutely simple, some base field extension of it would split into the direct sum of two ideals one of which would be the Jordan algebra of a pointed quadratic form of dimension 8 [R, Th. 1]. On the other hand, the property of being generated by two elements is inherited by this Jordan algebra, which by 7.2 is impossible. Thus $i_1 := \varphi \circ i$ and $i'_1 := \varphi' \circ i'$ satisfy all conditions of the lemma. □

Proposition 7.2.3. *Suppose J is split (which holds automatically if F is a finite field) and let $i\colon E \to J$ be an isomorphic embedding. Writing $K := \Delta(E)$ for the discriminant of E, there exists a subalgebra $J_1 \subseteq J$ such that*

$$i(E) \subseteq J_1 \cong \mathrm{Her}_3(K, \Gamma), \quad \Gamma := \mathrm{diag}(1, -1, 1).$$

Proof. Replacing E by $i(E)$ if necessary, we may assume $E \subseteq J$ and that $i\colon E \hookrightarrow J$ is the inclusion. We write $E^\perp \subseteq J$ for the orthogonal complement of E in J relative to the bilinear trace and, for all $v \in E^\perp$, denote by $q_E(v)$ the E-component of $v^\sharp$ relative to the decomposition $J = E \oplus E^\perp$. By [PeR 84a, Prop. 2.1], $E^\perp$ may be viewed as an E-module in a natural way, and $q_E\colon E^\perp \to E$ is a quadratic form over E. Moreover, combining [PeR 84a, Cor. 3.8] with a result of Engelberger [E, Prop. 1.2.5], we conclude that there exists an element $v \in E^\perp$ that is invertible in J and satisfies $q_E(v) = 0$. Now [PeR 84a, Prop. 2.2] yields a non-zero element $\alpha \in F$ such that the inclusion $E \hookrightarrow J$ can be extended to an isomorphic embedding from the étale first Tits construction $J(E, \alpha)$ into J. Write $J_1 \subseteq J$ for its image. Then $E \subseteq J_1 \cong J(E, \alpha)$, and from [PeR 84b, Th. 3] we deduce $J(E, \alpha) \cong \mathrm{Her}_3(K, \Gamma)$ with $\Gamma := \mathrm{diag}(1, -1, 1)$ as above. □

Proposition 7.2.4. *Let J_1, J'_1 be nine-dimensional absolutely simple subalgebras of J. Then every isotopy $J_1 \to J'_1$ can be extended to an element of the structure group of J.*

Proof. Let $\eta_1\colon J_1 \to J'_1$ be an isotopy. Then some $w \in J_1^\times$ makes $\eta_1\colon J_1^{(w)} \to J'_1$ an isomorphism. On the other hand, structure theory yields a central simple associative algebra (B, τ) of degree 3 with unitary involution over F and an isomorphism $\varphi\colon H(B, \tau) \to J_1$ which, setting $p := \varphi^{-1}(w) \in H(B, \tau)^\times$, may be regarded as an isomorphism

$$\varphi\colon H(B, \tau)^{(p)} \xrightarrow{\sim} J_1^{(w)}.$$

On the other hand, following (2.7.1),

$$R_p\colon H(B, \tau)^{(p)} \xrightarrow{\sim} H(B, \tau^{(p)})$$

is an isomorphism as well, and combining, we end up with an isomorphism

$$\varphi' := \eta_1 \circ \varphi \circ R_p^{-1} \colon\ H(B, \tau^{(p)}) \xrightarrow{\sim} J_1'.$$

Writing K for the centre of B and consulting 3.8, we now find a norm pair (u, μ) of (B, τ) such that φ extends to an isomorphism

$$\Phi\colon\ J(B, \tau, u, \mu) \xrightarrow{\sim} J.$$

Similarly, we find a norm pair (u', μ') of $(B, \tau^{(p)})$ such that φ' extends to an isomorphism

$$\Phi'\colon\ J(B, \tau^{(p)}, u', \mu') \xrightarrow{\sim} J.$$

Next, setting $u_1 := p^{\sharp -1}u'$, $\mu_1 := N_B(p)^{-1}\mu'$, we apply Lemma 3.8.5 to obtain an isotopy

$$\hat{R}_p\colon\ J(B, \tau, u_1, \mu_1) \to J(B, \tau^{(p)}, u', \mu'), \quad v_0 + vj \mapsto (v_0 p) + (p^{-1}vp)j, \tag{7.2.5}$$

and combining, we end up with an isotopy

$$\hat{R}_p^{-1} \circ \Phi'^{-1} \circ \Phi\colon\ J(B, \tau, u, \mu) \longrightarrow J(B, \tau, u_1, \mu_1).$$

Hence [Pe 04, Th. 5.2] yields an isomorphism

$$\Psi\colon\ J(B, \tau, u, \mu) \xrightarrow{\sim} J(B, \tau, u_1, \mu_1)$$

inducing the identity on $H(B, \tau)$. Thus

$$\eta := \Phi' \circ \hat{R}_p \circ \Psi \circ \Phi^{-1}\colon\ J \longrightarrow J$$

is an isotopy that fits into the diagram

$$\begin{tikzcd}
J(B,\tau,u_1,\mu_1) \arrow[dr, "\hat{R}_p"] & \\
J(B,\tau,u,\mu) \arrow[u, "\Psi"] \arrow[dddd, bend right=50, "\Phi"'] & J(B,\tau^{(p)},u',\mu') \arrow[dddd, bend left=50, "\Phi'"] \\
H(B,\tau) \arrow[u, hook] \arrow[r, "R_p"'] \arrow[d, "\varphi"'] & H(B,\tau^{(p)}) \arrow[u, hook] \arrow[d, "\varphi'"] \\
J_1 \arrow[r, "\eta_1"'] \arrow[d, hook] & J_1' \arrow[d, hook] \\
J \arrow[r, "\eta"'] & J,
\end{tikzcd}$$

whose arrows are either inclusions or isotopies. Now, since $\eta \circ \Phi = \Phi' \circ \hat{R}_p \circ \Psi$ by definition of η, and $\hat{R}_p$ agrees with R_p on $H(B, \tau)$ by (7.2.5), simple diagram chasing shows that $\eta \in \mathrm{Str}(J)$ is an extension of η_1. $\square$

We can now prove Theorem B in a form reminiscent of Th. 5.2.7.

Theorem 7.2.6. *Let J be an Albert algebra over F and E a cubic étale F-algebra. Then the pair (E, J) satisfies the weak Skolem-Noether property for isomorphic embeddings.*

Proof. Leit $i, i' : E \to J$ be two isomorphic embeddings. We must show that they are weakly equivalent and first claim that we may assume the following: *there exist a central simple associative algebra (B, τ) of degree 3 with unitary involution over F and a subalgebra $J_1 \subseteq J$ such that $J_1 \cong H(B, \tau)$ and i, i' factor uniquely through J_1 to isomorphic embeddings $i_1 : E \to J_1$, $i'_1 : E \to J_1$.* Indeed, replacing the isomorphic embeddings i, i' by strongly equivalent ones if necessary, this is clear by Lemma 7.2.2 provided F is infinite. On the other hand, if F is finite, Prop. 7.2.3 leads to absolutely simple nine-dimensional subalgebras $J_1, J'_1 \subseteq J$ that are isomorphic and have the property that i, i' factor uniquely through J_1, J'_1 to isomorphic embeddings $i_1 : E \to J_1$, $i'_1 : E \to J'_1$, respectively. But every isomorphism from J'_1 to J_1 extends to an automorphism of J [KMRT, 40.15], [Pe 04, Remark 5.6(b)]. Hence we may assume $J'_1 = J_1$, as desired.
With J_1, i_1, i'_1 as above, Th. 5.2.7 yields elements $w \in E$ of norm 1 and $\varphi_1 \in \mathrm{Str}(J_1)$ such that $i_1 \circ R_w = \varphi_1 \circ i'_1$. Using Prop. 7.2.4, we extend φ_1 to an element $\varphi \in \mathrm{Str}(J)$ and therefore conclude that the diagram (4.1.1) commutes. $\square$

8 Outer automorphisms for type ${}^3D_4$: proof of Theorem A

In this section, we apply Theorem B to prove Theorem A.

8.1 A subgroup of $\mathrm{Str}(J)$

Let E be a cubic étale subalgebra of an Albert algebra J and write H for the subgroup of $h \in \mathrm{Str}(J)$ that normalize E and such that $Nh = N$. Note that, for $\varphi \in \mathrm{Aut}(E)$, the element $\psi \in \mathrm{Str}(J)$ provided by Theorem B to extend $\varphi \circ R_w$ to all of J belongs to H. Indeed, as $\psi \in \mathrm{Str}(J)$, there is a $\mu \in F^\times$ such that $N\psi = \mu N$, but for $e \in E$ we have $N(\psi(e)) = N(\varphi(ew)) = N(\varphi(e))N(\varphi(w)) = N(e)$.
We now describe H in the case where J is a matrix Jordan algebra as in §3.5 with $\Gamma = \mathbf{1}_3$ and E is the subalgebra of diagonal matrices. We rely on some facts that are only proved in the literature under the hypothesis $\mathrm{char}\, F \neq 2, 3$. This hypothesis is not strictly necessary but we adopt it for now in order to ease the writing. Fix $h \in H$. The norm N restricts to E as $N(\sum \alpha_i e_{ii}) = \alpha_1\alpha_2\alpha_3$, so h permutes the three singular points $[e_{ii}]$ in the projective variety $N|_E = 0$ in $\mathbb{P}(E)$. There is an embedding of the symmetric group on 3 letters, Sym_3, in H acting by permuting the e_{ii} by their indices, see [Gar 06, §3.2] for an explicit formula, and consequently $H \cong H_0 \rtimes \mathrm{Sym}_3$, where H_0 is the subgroup of H of elements normalizing Fe_{ii} for each i. For $w := (w_1, w_2, w_3) \in (F^\times)^{\times 3}$ such that $w_1w_2w_3 = 1$, it follows that $U_w \in H$ (cf. (3.2.2)) sends $e_{ii} \mapsto w_i^2 e_{ii}$.

Assuming now that F is algebraically closed, after multiplying h by a suitable U_w, we may assume that h restricts to be the identity on E. The subgroup of such elements of $\mathrm{Str}(J)$ is identified with the $\mathrm{Spin}(C)$ which acts on the off-diagonal entries in J as a direct sum of the three inequivalent minuscule 8-dimensional representations, see [KMRT, 36.5, 38.6, 38.7] or [J 71, p. 18, Prop. 6]. Thus, we may identify H with $(R^{(1)}_{E/F}(\mathbb{G}_{\mathrm{m}}) \cdot \mathrm{Spin}(C)) \rtimes \mathrm{Sym}_3$, where Sym_3 acts via outer automorphisms on $\mathrm{Spin}(C)$ as in [Gar 06, §3] or [KMRT, 35.15].

8.2 The Tits class

Recall from §1.1 that the Dynkin diagram of a group G is endowed with an action by the absolute Galois group of F, and elements of $\mathrm{Aut}(\Delta)(F)$ act naturally on $H^2(F, Z)$, for Z the scheme-theoretic center of the simply connected cover of G.

Lemma 8.2.1. *Let G be a group of type D_4 over a field F with Dynkin diagram Δ. If there is a $\pi \in \mathrm{Aut}(\Delta)(F)$ of order 3 such that $\pi(t_G) = t_G$, then G has type ${}^1D_4$ or ${}^3D_4$ and $t_G = 0$.*

Proof. For the first claim, if G has type ${}^2D_4$ or ${}^6D_4$, then $\mathrm{Aut}(\Delta)(F) = \mathbb{Z}/2$ or 1.

Now suppose that G has type ${}^1D_4$. We may assume that G is simply connected. The center Z of G is $\mu_2 \times \mu_2$, with automorphism group Sym_3 and π acts on Z with order 3. The three nonzero characters $\chi_1, \chi_2, \chi_3 \colon Z \to \mathbb{G}_{\mathrm{m}}$ are permuted transitively by π, so by hypothesis the element $\chi_i(t_G) \in H^2(F, \mathbb{G}_{\mathrm{m}})$ does not depend on i. As the χ_i's satisfy the equations $\chi_1 + \chi_2 + \chi_3 = 0$ and $2\chi_i = 0$ (compare [T, 6.2] or [KMRT, 9.14]), it follows that $\chi_i(t_G) = 0$ for all i, hence $t_G = 0$ by [Gar 12, Prop. 7].

If G has type ${}^3D_4$, then there is a unique cyclic cubic field extension E of F such that $G \times E$ has type ${}^1D_4$. By the previous paragraph, restriction $H^2(F, Z) \to H^2(E, Z)$ kills t_G. That map is injective because Z has exponent 2, so $t_G = 0$. ☐

In the next result, the harder, "if" direction is the crux case of the proof of Theorem A and is an application of Theorem B. The easier, "only if" direction amounts to [CEKT, Th. 13.1] or [KT, Prop. 4.2]; we include it here as a consequence of the (a priori stronger) Lemma 8.2.1.

Proposition 8.2.2. *Let G be a group of type D_4 over a field F. The image of $\alpha(F) \colon \mathrm{Aut}(G)(F) \to \mathrm{Aut}(\Delta)(F)$ contains an element of order 3 if and only if G has type ${}^1D_4$ or ${}^3D_4$, G is simply connected or adjoint, and $t_G = 0$.*

Proof. *"If"* : We may assume that G is simply connected. If G has type ${}^1D_4$, then G is $\mathrm{Spin}(q)$ for some 3-Pfister quadratic form q, and the famous triality automorphisms of $\mathrm{Spin}(q)$ as in [SpV, 3.6.3, 3.6.4] are of order 3 and have image in $\mathrm{Aut}(\Delta)(F)$ of order 3. So assume G has type ${}^3D_4$.

Assume for this paragraph that char $F \neq 2, 3$. There is a uniquely determined cyclic Galois field extension E of F such that $G \times E$ has type ${}^1D_4$. By hypothesis, there is an Albert F-algebra J with norm form N such that $E \subset J$ and we may identify G with the algebraic group with K-points

$$\{g \in \mathrm{GL}(J \otimes K) \mid Ng = N \text{ and } g|_{E \otimes K} = \mathrm{Id}_{E \otimes K}\}$$

for every extension K of F. Take now φ to be a non-identity F-automorphism of E and $w \in E$ of norm 1 and $\psi \in \mathrm{Str}(J)$ to be the elements given by Theorem B such that $\psi|_E = \varphi \circ R_w$. As ψ normalizes E and preserves N, it follows immediately that ψ normalizes G as a subgroup of $\mathrm{Str}(J)$. (Alternatively this is obvious from the fact that in subsection 8.1, $\mathrm{Spin}(C)$ is the derived subgroup of H°.) Tracking through the description of H in subsection 8.1, we find that conjugation by ψ is an outer automorphism of G such that ψ^3 is inner.
In case F has characteristic 2 or 3, one can reduce to the case of characteristic zero as follows. Find R a complete discrete valuation ring with residue field F and fraction field K of characteristic zero. Lifting E to R allows us to construct a quasi-split simply connected group scheme $\mathcal{G}^q$ over R whose base change to F is the quasi-split inner form G^q of G. We have maps

$$H^1(F, G^q) \xleftarrow{\sim} H^1_{\text{ét}}(R, \mathcal{G}^q) \hookrightarrow H^1(K, \mathcal{G}^q \times K)$$

where the first map is an isomorphism by Hensel and the second map is injective by [BT]. Twisting by a well chosen $\mathcal{G}^q$-torsor, we obtain

$$H^1(F, G) \xleftarrow{\sim} H^1_{\text{ét}}(R, \mathcal{G}) \hookrightarrow H^1(K, \mathcal{G} \times K)$$

where $\mathcal{G} \times K$ has type ${}^3D_4$ and zero Tits class and $G \cong \mathcal{G} \times F$. Now in $\mathrm{Aut}(G)(F) \to \mathrm{Aut}(\Delta)(F) = \mathbb{Z}/3$, the inverse image of 1 is a connected component X of $\mathrm{Aut}(G)$ defined over F, a G-torsor. Lifting X to $H^1(K, \mathcal{G} \times K)$, we discover that this G-torsor is trivial (by the characteristic zero case of the theorem), hence X is F-trivial, i.e., has an F-point.

"Only if": Let $\phi \in \mathrm{Aut}(G)(F)$ be such that $\alpha(\phi)$ has order 3. In view of the inclusion (1.1.1), Lemma 8.2.1 applies. If G has type ${}^3D_4$, then it is necessarily simply connected or adjoint, so assume G has type ${}^1D_4$. Then ϕ lifts to an automorphism of the simply connected cover $\widetilde{G}$ of G, hence acts on the center Z of $\widetilde{G}$ in such a way that it preserves the kernel of the map $Z \to G$. As Z is isomorphic to $\mu_2 \times \mu_2$ and ϕ acts on it as an automorphism of order 3, the kernel must be 0 or Z, hence G is simply connected or adjoint. □

8.3 Proof of Theorem A

Let G be a group of type ${}^3D_4$, so $\mathrm{Aut}(\Delta)(F) = \mathbb{Z}/3$; put π for a generator. If $\pi(t_G) \neq t_G$, then the right side of (1.1.1) is a singleton and the containment is trivially an equality, so assume $\pi(t_G) = t_G$. Then $t_G = 0$ by Lemma 8.2.1 and the conclusion follows by Proposition 8.2.2. □

Example 8.3.1. Let F_0 be a field with a cubic Galois extension E_0. For the split adjoint group PSO_8 of type D_4 over F, a choice of pinning gives an isomorphism of $\mathrm{Aut}(\mathrm{PSO}_8)$ with $\mathrm{PSO}_8 \rtimes \mathrm{Sym}_3$, such that elements of Sym_3 normalize the Borel subgroup appearing in the pinning. Twisting Spin_8 by a 1-cocycle with values in $H^1(F_0, \mathrm{Sym}_3)$ representing the class of E_0 gives a simply connected quasi-split group G^q of type ${}^3D_4$. As in [GarMS, pp. 11, 12], there exists an extension F of F_0 and a versal torsor $\xi \in H^1(F, G^q)$; define G to be $G^q \times F$ twisted by ξ. As ξ is versal, the Rost invariant $r_{G^q}(\xi) \in H^3(F, \mathbb{Z}/6\mathbb{Z})$ has maximal order, namely 6 [GarMS, p. 149]. Moreover, the map $\alpha(F)\colon \mathrm{Aut}(G)(F) \to \mathrm{Aut}(\Delta)(F) = \mathbb{Z}/3$ is onto by Theorem A.
In case $\mathrm{char}\, F_0 \neq 2, 3$, G is $\mathrm{Aut}(\Gamma)$ for some twisted composition Γ in the sense of [KMRT, §36]. As $r_{G^q}(\xi)$ is not 2-torsion, by [KMRT, 40.16], Γ is not Hurwitz, and by [KT], $\mathrm{Aut}(G)(F)$ contains no outer automorphisms of order 3. This is a newly observed phenomenon, in that in all other cases where $\alpha(F)$ is known to be onto, it is also split.

9 Outer automorphisms for type A

9.1 Groups of type A_n

We now consider Conjecture 1.1.2 and Question 1.1.3 for groups G of type A_n. If G has inner type (i.e., is isogenous to $\mathrm{SL}_1(B)$ for a degree d central simple F-algebra) then equality holds in (1.1.1) and the answer to Question 1.1.3 is "yes" as in [Gar 12, p. 232].
So assume that G has outer type and in particular $n \geq 2$. The simply connected cover of G is $\mathrm{SU}(B, \tau)$ for B a central simple K-algebra of degree $d := n + 1$, where K is a quadratic étale F-algebra, and τ is a unitary K/F-involution. (This generalizes the (K, B, τ) defined in §3.7 by replacing 3 by d.) As the center Z of $\mathrm{SU}(B, \tau)$ is the group scheme $(\mu_d)_{[K]}$ of d-th roots of unity twisted by K in the sense of [KMRT, p. 418] (i.e., is the Cartier dual of the finite étale group scheme $(\mathbb{Z}/d)_{[K]}$), every subgroup of Z is characteristic, hence (1.1.1) is an equality for G if and only if it is so for $\mathrm{SU}(B, \tau)$ and similarly the answers to Question 1.1.3 are the same for G and $\mathrm{SU}(B, \tau)$. Therefore, we need only treat $\mathrm{SU}(B, \tau)$ below.
The automorphism group $\mathrm{Aut}(\Delta)(F)$ is $\mathbb{Z}/2$ and its nonzero element π acts on $H^2(F, Z)$ as -1, hence $\pi(t_{\mathrm{SU}(B,\tau)}) = -t_{\mathrm{SU}(B,\tau)}$ and the right side of (1.1.1) is a singleton (if $2t_{\mathrm{SU}(B,\tau)} \neq 0$) or has two elements (if $2t_{\mathrm{SU}(B,\tau)} = 0$). These cases are distinguished by the following lemma.

Lemma 9.1.1. *In case d is even (resp., odd): $2t_{\mathrm{SU}(B,\tau)} = 0$ if and only if $B \otimes_K B$ (resp., B) is a matrix algebra over K.*

Proof. The cocenter $Z^* := \mathrm{Hom}(Z, \mathbb{G}_{\mathrm{m}})$ is $(\mathbb{Z}/d)_{[K]}$; put $\chi_i \in Z^*$ for the element corresponding to $i \in (\mathbb{Z}/d)_{[K]}$. If $d = 2e$ for some integer e, then the element χ_e is fixed by $\mathrm{Gal}(F)$ and $2\chi_e = \chi_d = 0$, regardless of B or $t_{\mathrm{SU}(B,\tau)}$. All other χ_i have stabilizer subgroup $\mathrm{Gal}(K)$ and $\chi_i(2t_{\mathrm{SU}(B,\tau)}) \in H^2(K, Z)$

can be identified with the class of $B^{\otimes 2i}$ in the Brauer group of K, cf. [KMRT, p. 378].
The algebra $B \otimes_K B$ is a matrix algebra, then, if and only if χ_i vanishes on $2t_{\mathrm{SU}(B,\tau)}$ for all i. This is equivalent to $2t_{\mathrm{SU}(B,\tau)} = 0$ by [Gar 12, Prop. 7]. When the degree d of B is odd, $B \otimes_K B$ is a matrix algebra if and only if B is such. ☐

Corollary 9.1.2. *If G is a group of type A_n for n even, then equality holds in* (1.1.1) *and the answer to Question 1.1.3 is "yes".*

Proof. We may assume that G has outer type and is $\mathrm{SU}(B,\tau)$. If $2t_{\mathrm{SU}(B,\tau)} \neq 0$, then the right side of (1.1.1) is a singleton and the claim is trivial. Otherwise, by Lemma 9.1.1, B is a matrix algebra, i.e., $\mathrm{SU}(B,\tau)$ is the special unitary group of a K/F-hermitian form, and the claim follows. ☐

9.2

The algebraic group $\mathrm{Aut}(\mathrm{SU}(B,\tau))$ has two connected components: the identity component, which is identified with the adjoint group of $\mathrm{SU}(B,\tau)$, and the other component, whose F-points are the outer automorphisms of $\mathrm{SU}(B,\tau)$.

Theorem 9.2.1. *There is an isomorphism between the F-variety of K-linear anti-automorphisms of B commuting with τ and the non-identity component of* $\mathrm{Aut}\,\mathrm{SU}(B,\tau)$*, given by sending an anti-automorphism ψ to the outer automorphism $g \mapsto \psi(g)^{-1}$.*

Clearly, such an anti-automorphism provides an isomorphism of B with its opposite algebra, hence can only exist when B has exponent 2. This is a concrete illustration of the inclusion (1.1.1).

Proof. First suppose that F is separably closed, in which case we may identify $K = F \times F$, $B = M_d(F) \times M_d(F)$, and $\tau(b_1, b_2) = (b_2^t, b_1^t)$. A K-linear anti-automorphism ψ is, by Skolem-Noether, of the form $\psi(b_1, b_2) = (x_1 b_1^t x_1^{-1}, x_2 b_2^t x_2^{-1})$ for some $x_1, x_2 \in \mathrm{PGL}_d(F)$, and the assumption that $\psi\tau = \tau\psi$ forces that $x_2 = x_1^{-t}$.
As $\mathrm{Nrd}_{B/K}\,\psi = \mathrm{Nrd}_{B/K}$, it follows that ψ is an automorphism of the variety $\mathrm{SU}(B,\tau)$, hence ϕ defined by $\phi(g) := \psi(g)^{-1}$ is an automorphism of the group. As $\phi(b) = b^{-1}$ for $b \in K^\times$, i.e., ϕ acts nontrivially on the center, ϕ is an outer automorphism.
We have shown that there is a well-defined morphism from the variety of anti-automorphisms commuting with τ to the outer automorphisms of $\mathrm{SU}(B,\tau)$, and it remains to prove that it is an isomorphism. For this, note that PGL_d acts on $\mathrm{SU}(B,\tau)$ where the group action is just function composition, that this action is the natural action of the identity component of $\mathrm{SU}(B,\tau)$ on its other connected component, and that therefore the outer automorphisms are a PGL_d-torsor. Furthermore, the first paragraph of the proof showed that the anti-automorphisms commuting with τ also make up a PGL_d-torsor, where the

actions are related by $y.\psi = y^{-1}.\phi$ for $y \in \mathrm{PGL}_d$. This completes the proof for F separably closed.

For general F, we note that the map $\psi \mapsto \phi$ is F-defined and gives an isomorphism over F_{sep}, hence is an isomorphism over F. ☐

9.3

We do not know how to prove or disprove existence of an anti-automorphism commuting with τ in general, but we can give a criterion for Question 1.1.3 that is analogous to the one given in [KT] for groups of type ${}^3D_4$.

Corollary 9.3.1. *A group* $\mathrm{SU}(B,\tau)$ *of outer type A has an F-defined outer automorphism of order 2 if and only if there exists a central simple algebra* (B_0,τ_0) *over F with* τ_0 *an involution of the first kind such that* (B,τ) *is isomorphic to* $(B_0 \otimes K, \tau_0 \otimes \iota)$*, for* ι *the non-identity F-automorphism of K.*

Proof. The bijection in Theorem 9.2.1 identifies outer automorphisms of order 2 with anti-automorphisms of order 2. If such a (B_0,τ_0) exists, then clearly τ_0 provides an anti-automorphism of order 2.

Conversely, given an anti-automorphism τ_0 of order 2, define a semilinear automorphism of B via $\iota := \tau_0\tau$. Set $B_0 := \{b \in B \mid \iota(b) = b\}$; it is an F-subalgebra and τ_0 restricts to be an involution on B_0. ☐

Example 9.3.2. We now exhibit a (B,τ) with B of exponent 2, but such that $\mathrm{SU}(B,\tau)$ has no outer automorphism of order 2 over F. The paper [ART] provides a field F and a division F-algebra C of degree 8 and exponent 2 such that C is not a tensor product of quaternion algebras. Moreover, it provides a quadratic extension K/F contained in C. It follows that $C \otimes K$ has index 4, and we set B to be the underlying division algebra. As $\mathrm{cor}_{K/F}[B] = 2[C] = 0$ in the Brauer group, B has a unitary involution τ.

For sake of contradiction, suppose that $\mathrm{SU}(B,\tau)$ had an outer automorphism of order 2, hence there exists a (B_0,τ_0) as in Corollary 9.3.1. Then B_0 has degree 4, so B_0 is a biquaternion algebra. Moreover, $C \otimes B_0$ is split by K, hence is Brauer-equivalent to a quaternion algebra Q. By comparing degrees, we deduce that C is isomorphic to $B_0 \otimes Q$, contradicting the choice of C.

9.4 Type ${}^2E_6$

Results entirely analogous to Theorem 9.2.1, Corollary 9.3.1, and Example 9.3.2 also hold for groups G of type ${}^2E_6$, using proofs of a similar flavor. The Dynkin diagram of type E_6 has automorphism group $\mathbb{Z}/2 = \{\mathrm{Id}, \pi\}$, and arguing as in Lemmas 8.2.1 or 9.1.1 shows that $\pi(t_G) = t_G$ if and only if $t_G = 0$. So for addressing Conjecture 1.1.2 and Question 1.1.3, it suffices to consider only those groups with zero Tits class, which can be completely described in terms of the hermitian Jordan triples introduced in [GarP, §4] or the Brown algebras studied in [Gar 01]. We leave the details to the interested reader.

Does Conjecture 1.1.2 hold for every group of type ${}^2E_6$? One might hope to imitate the outline of the proof of Theorem A. Does an analogue of Theorem B hold, where one replaces Albert algebras, cubic Galois extensions, and the inclusion of root systems $D_4 \subset E_6$ by Brown algebras or Freudenthal triple systems, quadratic Galois extensions, and the inclusion $E_6 \subset E_7$?

Acknowledgements

The first author was partially supported by NSF grant DMS-1201542. He thanks TU Dortmund for its hospitality under the auspices of the Gambrinus Fellowship while some of the research was conducted. The second author thanks the Department of Mathematics and Statistics of the University of Ottawa, where much of this research was initiated, for its support and hospitality.

Both authors would like to express their gratitude to the referees, who read an earlier version of the paper with great care and made numerous useful suggestions for improvement.

References

[AJ] A.A. Albert and N. Jacobson, *On reduced exceptional simple Jordan algebras*, Ann. of Math. (2) 66 (1957), 400–417. MR 0088487 (19,527b)

[ART] S.A. Amitsur, L.H. Rowen, and J.-P. Tignol, *Division algebras of degree* 4 *and* 8 *with involution*, Israel J. Math. 33 (1979), 133–148.

[Br] M. Brion, *On extensions of algebraic groups with finite quotients*, Pacific J. Math. 279 (2015), no. 1–2, 135–153.

[BT] F. Bruhat and J. Tits, *Groupes algébriques sur un corps local. Chapitre III. Compléments et applications á la cohomologie galoisienne*, J. Fac. Sci. Univ. Tokyo Sect. IA Math. 34 (1987), no. 3, 671–698.

[CEKT] V. Chernousov, A. Elduque, M-A. Knus, and J-P. Tignol, *Algebraic groups of type D_4, triality, and composition algebras*, Documenta Math. 18 (2013), 413–468.

[CKT] V. Chernousov, M-A. Knus, and J-P. Tignol, *Conjugacy class of trialitarian automorphisms and symmetric compositions*, J. Ramanujan Math. Soc. 27 (2012), 479–508.

[DG] M. Demazure and A. Grothendieck, *SGA3: Schémas en groupes*, Lecture Notes in Mathematics, vol. 151–153, Springer, 1970.

[E] R. Engelberger, *Twisted compositions*, Ph.D. thesis, ETH Zürich, 2002.

[GanS] W.T. Gan and G. Savin, *Twisted Bhargava cubes*, Algebra & Number Theory 8 (2014), no. 8, 1913–1957.

[Gar 01] S. Garibaldi, *Structurable algebras and groups of type E_6 and E_7*, J. Algebra 236 (2001), no. 2, 651–691.

[Gar 06] ______, *Unramified cohomology of classifying varieties for exceptional simply connected groups*, Trans. Amer. Math. Soc. 358 (2006), no. 1, 359–371.

[Gar 12] ______, *Outer automorphisms of algebraic groups and determining groups by their maximal tori*, Michigan Math. J. 61 (2012), no. 2, 227–237.

[GarMS] S. Garibaldi, A. Merkurjev, and J-P. Serre, *Cohomological invariants in Galois cohomology*, University Lecture Series, vol. 28, Amer. Math. Soc., 2003.

[GarP] S. Garibaldi and H.P. Petersson, *Groups of outer type E_6 with trivial Tits algebras*, Transf. Groups 12 (2007), no. 3, 443–474.

[J 71] N. Jacobson, *Exceptional Lie algebras*, Lecture notes in pure and applied mathematics, vol. 1, Marcel-Dekker, New York, 1971.

[J 81] N. Jacobson, *Structure theory of Jordan algebras*, University of Arkansas Lecture Notes in Mathematics, vol. 5, University of Arkansas, Fayetteville, Ark., 1981. MR 634508 (83b:17015)

[JK] N. Jacobson and J. Katz, *Generically algebraic quadratic Jordan algebras*, Scripta Math. 29 (1973), no. 3-4, 215–227, Collection of articles dedicated to the memory of Abraham Adrian Albert. MR 0404370 (53 #8172)

[KMRT] M.-A. Knus, A.S. Merkurjev, M. Rost, and J.-P. Tignol, *The book of involutions*, Colloquium Publications, vol. 44, Amer. Math. Soc., 1998.

[KT] M-A. Knus and J-P. Tignol, *Triality and algebraic groups of type ${}^3D_4$*, Doc. Math. (2015), 387–405, Extra Volume: Alexander S. Merkurjev's Sixtieth Birthday.

[Lang] S. Lang, *Algebra*, third ed., Graduate Texts in Mathematics, vol. 211, Springer-Verlag, New York, 2002. MR 1878556 (2003e:00003)

[Loos] O. Loos, *Jordan pairs*, Lecture Notes in Mathematics, vol. 460, Springer-Verlag, Berlin, 1975.

[McC 69a] K. McCrimmon, *The Freudenthal-Springer-Tits constructions of exceptional Jordan algebras*, Trans. Amer. Math. Soc. 139 (1969), 495–510. MR 0238916 (39 #276)

[McC 69b] ______, *On Herstein's theorems relating Jordan and associative algebras*, J. Algebra 13 (1969), 382–392.

[McC 70] ______, *The Freudenthal-Springer-Tits constructions revisited*, Trans. Amer. Math. Soc. 148 (1970), 293–314.

[McCZ] K. McCrimmon and E. Zelmanov, *The structure of strongly prime quadratic Jordan algebras*, Adv. in Math. 69 (1988), no. 2, 133–222. MR 946263 (89k:17052)

[MT] A.S. Merkurjev and J.-P. Tignol, *The multipliers of similitudes and the Brauer group of homogeneous varieties*, J. reine angew. Math. 461 (1995), 13–47.

[Pe 78] H.P. Petersson, *Conjugacy of idempotents in Jordan pairs*, Comm. Algebra 6 (1978), no. 7, 673–715.

[Pe 04] ———, *Structure theorems for Jordan algebras of degree three over fields of arbitrary characteristic*, Comm. Algebra 32 (2004), no. 3, 1019–1049.

[Pe 15] ———, *An embedding theorem for reduced Albert algebras over arbitrary fields*, Comm. Algebra 43 (2015), no. 5, 2062–2088.

[PeR 84a] H.P. Petersson and M.L. Racine, *Springer forms and the first Tits construction of exceptional Jordan division algebras*, Manuscripta Math. 45 (1984), no. 3, 249–272. MR 85i:17029

[PeR 84b] ———, *The toral Tits process of Jordan algebras*, Abh. Math. Sem. Univ. Hamburg 54 (1984), 251–256. MR 86g:17020

[PeT04a] H.P. Petersson and M. Thakur, *The étale Tits process of Jordan algebras revisited*, J. Algebra 273 (2004), no. 1, 88–107. MR 2032452 (2004m:17047)

[PrT] R. Preeti and J.-P. Tignol, *Multipliers of improper similitudes*, Documenta Math. 9 (2004), 183–204.

[R] M.L. Racine, *A note on quadratic Jordan algebras of degree* 3, Trans. Amer. Math. Soc. 164 (1972), 93–103. MR 0304447 (46 #3582)

[Se] J-P. Serre, *Gèbres*, L'Enseignement Math. 39 (1993), 33–85, (reprinted in Oeuvres, vol. IV, pp. 272–324).

[Sp] T.A. Springer, *Linear algebraic groups*, second ed., Birkhäuser, 1998.

[SpV] T.A. Springer and F.D. Veldkamp, *On Hjelmslev-Moufang planes*, Math. Zeit. 107 (1968), 249–263.

[T] J. Tits, *Représentations linéaires irréductibles d'un groupe réductif sur un corps quelconque*, J. Reine Angew. Math. 247 (1971), 196–220.

Skip Garibaldi
UCLA Institute for
Pure & Applied Mathematics
460 Portola Plaza
Box 957121
Los Angeles, CA 90095-7121
USA

Holger P. Petersson
Fakultät für Mathematik
und Informatik
FernUniversität in Hagen
D-58084 Hagen
Germany

Gerstenhaber-Schack and Hochschild Cohomologies of Hopf Algebras

Julien Bichon

Received: May 21, 2015
Revised: December 17, 2015

Communicated by Joachim Cuntz

Abstract. We show that the Gerstenhaber-Schack cohomology of a Hopf algebra determines its Hochschild cohomology, and in particular its Gerstenhaber-Schack cohomological dimension bounds its Hochschild cohomological dimension, with equality of the dimensions when the Hopf algebra is cosemisimple of Kac type. Together with some general considerations on free Yetter-Drinfeld modules over adjoint Hopf subalgebras and the monoidal invariance of Gerstenhaber-Schack cohomology, this is used to show that both Gerstenhaber-Schack and Hochschild cohomological dimensions of the coordinate algebra of the quantum permutation group are 3.

2010 Mathematics Subject Classification: 16T05, 16E40, 16E10
Keywords and Phrases: Hopf algebra, cohomology, cohomological dimension, Yetter-Drinfeld module

1 Introduction

We study homological properties of Hopf algebras by using Yetter-Drinfeld modules and tensor category techniques. We are especially interested in the following question:

Question 1.1. *If A and B are Hopf algebras having equivalent tensor categories of comodules, how are their Hochschild cohomologies related? In particular do A and B have the same cohomological dimension?*

We have seen in [10] that the Hochschild cohomologies of two such Hopf algebras A and B are indeed closely related, using resolutions of the trivial Yetter-Drinfeld module over A (or over B) formed by free Yetter-Drinfeld modules. In the present paper we continue this study along the same line of ideas.

Our first remark in view of Question 1.1 is that there exists at least a cohomology theory for Hopf algebras that is known to be well-behaved with respect to this situation: Gerstenhaber-Schack cohomology [27, 28]. Let A be a Hopf algebra and let M be a Hopf bimodule over A: the Gerstenhaber-Schack cohomology $H^*_{\mathrm{GS}}(A, M)$ of A with coefficients in M [28] is defined to be the homology of an explicit bicomplex whose columns are modeled on the Hochschild complex of the underlying algebra and rows are modeled on the Cartier complex of the underlying coalgebra. When $M = A$ is the trivial Hopf bimodule, then $H^*_{\mathrm{GS}}(A, A) =: H^*_b(A)$ is known as the bialgebra cohomology of A. This cohomology theory, which can also be defined in terms of Yetter-Drinfeld modules, was first introduced in view of applications to deformation theory, and has been used as a key tool in the proof of several fundamental results on finite-dimensional Hopf algebras [55, 23].

If A and B are Hopf algebras having equivalent tensor categories of comodules, then there exists a tensor equivalence $F : {}^A_A\mathcal{M}^A_A \to {}^B_B\mathcal{M}^B_B$ between their categories of Hopf bimodules such that for any Hopf bimodule M over A, we have $H^*_{\mathrm{GS}}(A, M) \simeq H^*_{\mathrm{GS}}(B, F(M))$, and in particular $H^*_b(A) \simeq H^*_b(B)$ and $\mathrm{cd}_{\mathrm{GS}}(A) = \mathrm{cd}_{\mathrm{GS}}(B)$ (where $\mathrm{cd}_{\mathrm{GS}}$ denotes the Gerstenhaber-Schack cohomological dimension, defined in the obvious way, see Section 5). We call these properties the monoidal invariance of Gerstenhaber-Schack cohomology.

Going back to Question 1.1, the next question is to study how Hochschild and Gerstenhaber-Schack cohomologies are related. We show that the Gerstenhaber-Schack cohomology of a Hopf algebra determines its Hochschild cohomology. More precisely, we show that if A is a Hopf algebra, then there exists a functor $G : {}_A\mathcal{M}_A \to {}^A_A\mathcal{M}^A_A$ such that for any A-bimodule M, we have

$$H^*(A, M) \simeq H^*_{\mathrm{GS}}(A, G(M))$$

In particular we have $\mathrm{cd}(A) \leq \mathrm{cd}_{\mathrm{GS}}(A)$. Then if A and B are Hopf algebras as in Question 1.1, combining this with the monoidal invariance of Gerstenhaber-Schack cohomology, we get the existence of two functors $F_1 : {}_A\mathcal{M}_A \to {}^B_B\mathcal{M}^B_B$ and $F_2 : {}_B\mathcal{M}_B \to {}^A_A\mathcal{M}^A_A$ such that for any A-bimodule M and any B-bimodule N, we have

$$H^*(A, M) \simeq H^*_{\mathrm{GS}}(B, F_1(M)) \text{ and } H^*(B, N) \simeq H^*_{\mathrm{GS}}(A, F_2(N))$$

In particular

$$\max(\mathrm{cd}(A), \mathrm{cd}(B)) \leq \mathrm{cd}_{\mathrm{GS}}(A) = \mathrm{cd}_{\mathrm{GS}}(B)$$

These isomorphisms and this inequality thus provide partial answers to Question 1.1. They lead to the following new question:

Question 1.2. *Is it true that* $\mathrm{cd}(A) = \mathrm{cd}_{\mathrm{GS}}(A)$ *for any Hopf algebra* A *over a field of characteristic zero? Is it true at least if* A *is assumed to be cosemisimple?*

A positive answer would give the monoidal invariance of cohomological dimension and fully answer the last part of Question 1.1, and would also be a natural infinite-dimensional generalization of a famous result by Larson-Radford [38], which states that, in characteristic zero, a finite-dimensional cosemisimple Hopf algebra is semisimple. See Remark 5.8.

We provide (Corollary 5.10) a partial positive answer to Question 1.2 in the case where A is cosemisimple of Kac type (the square of the antipode is the identity), and in turn this gives a partial positive answer to Question 1.1 (see Corollary 5.11).

We then apply our general considerations to quantum symmetry Hopf algebras, which were the first motivation for this work. Let (R, φ) semisimple measured algebra of dimension ≥ 4, and let $A_{\mathrm{aut}}(R, \varphi)$ be its quantum symmetry Hopf algebra [61, 8]. We compute, in the cosemisimple case, the bialgebra cohomology of $A_{\mathrm{aut}}(R, \varphi)$, and we show that $\mathrm{cd}(A_{\mathrm{aut}}(R, \varphi)) \leq \mathrm{cd}_{\mathrm{GS}}(A_{\mathrm{aut}}(R, \varphi)) = 3$, with equality if φ is a trace. These results include in particular the coordinate algebra of Wang's quantum permutation group S_n^+ [61].

As a last comment to further motivate the use of Gerstenhaber-Schack cohomology as an appropriate cohomology theory for Hopf algebras (apart from its use to get information on Hochschild cohomology itself), we would like to point out that, in the examples computed so far, it also has the merit to avoid the "dimension drop" phenomenon usually encountered for quantum algebras (see [31, 32]): the canonical choice of coefficients (the trivial Hopf bimodule) is the good one to get the cohomological dimension. It would be interesting to know if this can be further generalized.

The paper is organized as follows. Section 2 consists of preliminaries. In Section 3 we discuss the cohomological dimension of a Hopf subalgebra and the sub-additivity of the cohomological dimension under extensions. Section 4 is devoted to Yetter-Drinfeld modules: we recall the concept of free (resp. co-free) Yetter Drinfeld module and we introduce the notion of relative projective (resp. injective) Yetter-Drinfeld module, which corresponds, via the tensor equivalence between Yetter-Drinfeld modules and Hopf bimodules [49], to the notion of relative projective (resp. injective) Hopf bimodule considered in [53]. We show that relative projective (resp. injective) Yetter-Drinfeld modules are precisely the direct summands of free (resp. co-free) Yetter-Drinfeld modules. This section also contains some considerations on free Yetter-Drinfeld modules over adjoint Hopf subalgebras. In Section 5, after having recalled some basic facts on Gerstenhaber-Schack cohomology, we provide an explicit complex that computes the Gerstenhaber-Schack cohomology $H^*_{\mathrm{GS}}(A, V)$, if A is cosemisimple or if the Yetter-Drinfeld module V is relative injective, using results from [53] in this last case (see Proposition 5.3). We then show that Gerstenhaber-Schack cohomology determines Hochschild cohomology, and show that Question 1.2 has a positive answer in the case of cosemisimple Hopf algebras of Kac type. In Section 6 we study the examples mentioned earlier in the introduction. In Section 7 we discuss the Gerstenhaber-Schack cohomological dimension in the setting of Hopf algebras having a projection.

2 Preliminaries

In this preliminary section we fix some notation, we recall some basic definitions and facts on the Hochschild cohomology of a Hopf algebra, and we discuss exact sequences of Hopf algebras.

2.1 Notations and conventions

We work over $\mathbb{C}$ (or over any algebraically closed field of characteristic zero). This assumption does not affect any of the theoretical results in the paper, but is important for the examples we consider. We assume that the reader is familiar with the theory of Hopf algebras and their tensor categories of comodules, as e.g. in [34, 35, 42]. If A is a Hopf algebra, as usual, Δ, ε and S stand respectively for the comultiplication, counit and antipode of A. We use Sweedler's notations in the standard way. The category of right A-comodules is denoted $\mathcal{M}^A$, the category of right A-modules is denoted $\mathcal{M}_A$, etc... The trivial (right) A-module is denoted $\mathbb{C}_\varepsilon$. The set of A-module morphisms (resp. A-comodule morphisms) between two A-modules (resp. two A-comodules) V and W is denoted $\mathrm{Hom}_A(V,W)$ (resp. $\mathrm{Hom}^A(V,W)$).

2.2 Hochschild cohomology of a Hopf algebra

If A is an algebra and M is an A-bimodule, then $H^*(A,M)$ denotes, as usual, the Hochschild cohomology of A with coefficients in M. See e.g. [62].

Definition 2.1. *The Hochschild cohomological dimension of an algebra A is defined to be*

$$\mathrm{cd}(A) = \sup\{n : \ H^n(A,M) \neq 0 \text{ for some } A-\text{bimodule } M\} \in \mathbb{N} \cup \{\infty\}$$

As noted by several authors (see [25], [29], [31], [13], [17], [10]), the Hochschild cohomology of a Hopf algebra can be described by using a suitable Ext functor on the category of left or right A-modules. Indeed, if A is a Hopf algebra and M is an A-bimodule, we have

$$H^*(A,M) \simeq \mathrm{Ext}^*_A(\mathbb{C}_\varepsilon, M')$$

where the above Ext is in the category of right A-modules and M' is M equipped with the right A-module structure given by $x \leftarrow a = S(a_{(1)}) \cdot x \cdot a_{(2)}$. This leads to the following description of the cohomological dimension of a Hopf algebra.

Proposition 2.2. *Let A be a Hopf algebra. We have*

$$\begin{aligned}\mathrm{cd}(A) &= \sup\{n : \ \mathrm{Ext}^n_A(\mathbb{C}_\varepsilon, M) \neq 0 \text{ for some } A-\text{module } M\}\\ &= \inf\{n : \ \mathrm{Ext}^i_A(\mathbb{C}_\varepsilon, -) = 0 \text{ for } i > n\}\\ &= \inf\{n : \ \mathbb{C}_\varepsilon \text{ admits a projective resolution of length } n\}\end{aligned}$$

Proof. The previous isomorphism ensures that

$$\mathrm{cd}(A) \leq \sup\{n : \ \mathrm{Ext}^n_A(\mathbb{C}_\varepsilon, M) \neq 0 \text{ for some } A-\text{module } M\}$$

If V is a right A-module, let ${}_\varepsilon V$ be the A-bimodule whose right structure is that of V and whose left structure is trivial, i.e. given by ε. Then $({}_\varepsilon V)' = V$, hence the converse inequality holds, and the first equality in the statement is proved, as well as the second one. The last one is shown similarly as in the case of group cohomology, see e.g. [14, Chapter VIII, Lemma 2.1]. □

Examples 2.3.

1. If G is a linear algebraic group, with coordinate algebra $\mathcal{O}(G)$, it is well-known that $\mathrm{cd}(\mathcal{O}(G)) = \dim G$.
2. If Γ is a (discrete) group, then $\mathrm{cd}(\mathbb{C}\Gamma) = \mathrm{cd}_{\mathbb{C}}(\Gamma)$, the cohomological dimension of Γ with coefficients $\mathbb{C}$. We have $\mathrm{cd}(\mathbb{C}\Gamma) = 0$ if and only if Γ is finite, and if Γ is finitely generated, then $\mathrm{cd}(\mathbb{C}\Gamma) = 1$ if and only if Γ contains a free normal subgroup of finite index, see [22].
3. If A is a finite-dimensional Hopf algebra, then either $\mathrm{cd}(A) = 0$ (when A is semisimple) or $\mathrm{cd}(A) = \infty$, a finite-dimensional Hopf algebra being Frobenius and hence self-injective.

2.3 Exact sequences of Hopf algebras

A sequence of Hopf algebra maps

$$\mathbb{C} \to B \xrightarrow{i} A \xrightarrow{p} L \to \mathbb{C}$$

is said to be exact [3] if the following conditions hold:

1. i is injective and p is surjective,
2. $\ker p = Ai(B)^+ = i(B)^+A$, where $i(B)^+ = i(B) \cap \mathrm{Ker}(\varepsilon)$,
3. $i(B) = A^{\mathrm{co}\, L} = \{a \in A : (\mathrm{id} \otimes p)\Delta(a) = a \otimes 1\} = {}^{\mathrm{co}\, L}A = \{a \in A : (p \otimes \mathrm{id})\Delta(a) = 1 \otimes a\}$.

Note that condition (2) implies $pi = \varepsilon 1$.

Proposition 2.4. *Let*

$$\mathbb{C} \to B \xrightarrow{i} A \xrightarrow{p} L \to \mathbb{C}$$

be a sequence of Hopf algebra maps where i is injective, p is surjective and $pi = \varepsilon 1$. Assume that the antipode of A is bijective. Consider the following three assertions.

1. *A is faithfully flat as a right B-module and* $\mathrm{Ker}(p) = Ai(B)^+ = i(B)^+A$.

2. ${}^{\mathrm{co}\,L}A = A^{\mathrm{co}\,L} = i(B)$ *and p is left or right faithfully coflat.*

3. *The sequence is exact.*

Then we have (1) $\Rightarrow$ (3) *and* (2) $\Rightarrow$ (3)*, and if* (3) *holds, then we have* (1) $\iff$ (2)*.*

An exact sequence satisfying (1) and (2) is called strict [52]. Note that p is automatically faithfully coflat if L is cosemisimple.
That (1) $\Rightarrow$ (3) holds is well-known (see [3, Proposition 1.2.4], [51, Lemma 1.3], [42, Proposition 3.4.3], or more generally [59, Theorem 1]). Also that (1) $\iff$ (2) if (3) holds is known, see [51, Corollary 1.8]. I wish to thank the referee for pointing out that (2) $\Rightarrow$ (3) follows from [59, Theorem 2], combined with [44, Remark 1.3].

3 Cohomological dimension of a Hopf subalgebra

In this section we discuss the behavior of cohomological dimension when passing to a Hopf subalgebra, which, under mild assumptions, is similar to the group cohomology case.

Proposition 3.1. *Let $B \subset A$ be a Hopf subalgebra. Assume that one of the following conditions holds.*

1. *A is projective as a right B-module.*

2. *The antipode of A is bijective and A is faithfully flat as a right B-module.*

3. *A is cosemisimple.*

4. *The exists a Hopf algebra map $\pi : A \to B$ such that $\pi_{|B} = \mathrm{id}_B$.*

5. *The antipode of A is bijective and B is commutative.*

Then $\mathrm{cd}(B) \leq \mathrm{cd}(A)$*.*

Proof. If A is projective as a right B-module, any projective right A-module is projective as a right B-module, thus a resolution of length n of $\mathbb{C}_\varepsilon$ in $\mathcal{M}_A$ yields a resolution of length n in $\mathcal{M}_B$, and thus Proposition 2.2 ensures that $\mathrm{cd}(B) \leq \mathrm{cd}(A)$. Assuming (2), Corollary 1.8 in [51] yields that A is projective as a right B-module, and we conclude by (1). If we assume that A is cosemisimple, then its antipode is bijective and by [16] A is faithfully flat as a right B-module, and we conclude by (2). If we assume (4), then A is free as a right B-module, see [47] (we will come back to this situation in Section 7), thus we conclude by (1). If B is commutative, then A is faithfully flat over B by Proposition 3.12 in [4], and again we conclude by (2). $\square$

The following result is the generalization of the sub-additivity of cohomological dimension under extensions (see e.g. Proposition 2.4 in [14]) with essentially the same proof, using Stefan's spectral sequence [54] as the natural generalization of the Hochschild-Serre spectral sequence.

Proposition 3.2. *Let*

$$\mathbb{C} \longrightarrow B \xrightarrow{i} A \xrightarrow{p} L \longrightarrow \mathbb{C}$$

be a strict exact sequence of Hopf algebras , and assume that the antipode of A is bijective. Then we have $\mathrm{cd}(B) \leq \mathrm{cd}(A) \leq \mathrm{cd}(B) + \mathrm{cd}(L)$. *If moreover L is semisimple, then* $\mathrm{cd}(B) = \mathrm{cd}(A)$.

Proof. By [51, Lemma 1.3], (or more generally [59, Theorem 1], see also [42, Proposition 3.4.3]), the canonical map

$$\begin{aligned} A \otimes_B A &\longrightarrow A \otimes L \\ a \otimes_B a' &\longmapsto aa'_{(1)} \otimes p(a'_{(2)}) \end{aligned}$$

is bijective. Thus $B \subset A$ is an L-Galois extension, and A is faithfully flat both as a left and right B-module (the antipode of A is bijective). Thus for any A-A-bimodule M there exists a spectral sequence [54]

$$E_2^{pq} = H^p(L; H^q(B, M)) \Rightarrow H^{p+q}(A, M)$$

The spectral sequence is concentrated in the rectangle $0 \leq p \leq \mathrm{cd}(L)$, $0 \leq q \leq \mathrm{cd}(B)$, and it follows that for $i > \mathrm{cd}(L) + \mathrm{cd}(B)$, we have $H^i(A, M) = 0$, and this proves the inequality. If L is semisimple, then $\mathrm{cd}(L) = 0$, and hence $\mathrm{cd}(B) = \mathrm{cd}(A)$. $\square$

4 Yetter-Drinfeld modules

Let A be a Hopf algebra. Recall that a (right-right) Yetter-Drinfeld module over A is a right A-comodule and right A-module V satisfying the condition, $\forall v \in V$, $\forall a \in A$,

$$(v \leftarrow a)_{(0)} \otimes (v \leftarrow a)_{(1)} = v_{(0)} \leftarrow a_{(2)} \otimes S(a_{(1)})v_{(1)}a_{(3)}$$

The category of Yetter-Drinfeld modules over A is denoted $\mathcal{YD}_A^A$: the morphisms are the A-linear A-colinear maps. Endowed with the usual tensor product of modules and comodules, it is a tensor category, with unit the trivial Yetter-Drinfeld module, denoted $\mathbb{C}$.

An important example of Yetter-Drinfeld module is the right coadjoint Yetter-Drinfeld module A_{coad}: as a right A-module $A_{\mathrm{coad}} = A$ and the right A-comodule structure is defined by

$$\mathrm{ad}_r(a) = a_{(2)} \otimes S(a_{(1)})a_{(3)}, \forall a \in A$$

The coadjoint Yetter-Drinfeld module has a natural generalization, discussed in the next subsection.

4.1 Free and co-free Yetter-Drinfeld modules

We now discuss some important constructions of Yetter-Drinfeld modules (left-right versions of these constructions were first given in [15], see [53] as well, in the context of Hopf bimodules).
Let V be a right A-comodule. The Yetter-Drinfeld module $V \boxtimes A$ is defined as follows [10]. As a vector space $V \boxtimes A = V \otimes A$, the right module structure is given by multiplication on the right, and the right coaction $\alpha_{V\boxtimes A}$ is defined by

$$\alpha_{V\boxtimes A}(v \otimes a) = v_{(0)} \otimes a_{(2)} \otimes S(a_{(1)})v_{(1)}a_{(3)}$$

Note that $A_{\mathrm{coad}} = \mathbb{C} \boxtimes A$.
A Yetter-Drinfeld module is said to be free if it is isomorphic to $V \boxtimes A$ for some comodule V.
The construction of the free Yetter-Drinfeld module on a comodule yields a functor $L = - \boxtimes A : \mathcal{M}^A \longrightarrow \mathcal{YD}_A^A$ which is left adjoint to the forgetful functor $R : \mathcal{YD}_A^A \longrightarrow \mathcal{M}^A$. Indeed we have natural isomorphisms

$$\begin{aligned} \mathrm{Hom}^A(V, R(X)) &\longrightarrow \mathrm{Hom}_{\mathcal{YD}_A^A}(V \boxtimes A, X) \\ f &\longmapsto \tilde{f},\ \tilde{f}(v \otimes a) = f(v) \leftarrow a \end{aligned}$$

for any A-comodule V and any Yetter-Drinfeld module X.
Now let M be a right A-module. The Yetter-drinfeld module $M\#A$ is defined as follows: the underlying vector space is $M \otimes A$, the right coaction is $\mathrm{id}_M \otimes \Delta$, while the right action is given by

$$(x \otimes a) \leftarrow b = x \cdot b_{(2)} \otimes S(b_{(1)})ab_{(3)}$$

The Yetter-Drinfeld module $\mathbb{C}\#A$ is the adjoint Yetter-Drinfeld module, denoted A_{ad}.
A Yetter-Drinfeld module will be said to be co-free if it is isomorphic to $M\#A$ for some module M. The construction of the co-free Yetter-Drinfeld module on a module yields a functor $L = -\#A : \mathcal{M}_A \longrightarrow \mathcal{YD}_A^A$ which is right adjoint to the forgetful functor $L : \mathcal{YD}_A^A \longrightarrow \mathcal{M}_A$. Indeed we have natural isomorphisms

$$\begin{aligned} \mathrm{Hom}_{\mathcal{YD}_A^A}(X, M\#A) &\longrightarrow \mathrm{Hom}_A(L(X), M) \\ f &\longmapsto (\mathrm{id}_M \otimes \varepsilon)f, \end{aligned}$$

for any A-module M and any Yetter-Drinfeld module X.

4.2 Relative projective and relative injective Yetter-Drinfeld modules

We will use the following notions.

Definition 4.1. *Let V be a Yetter-Drinfeld module over A. Then V is said to be relative projective if the functor $\mathrm{Hom}_{\mathcal{YD}_A^A}(V, -)$ transforms exact sequences*

of Yetter-Drinfeld modules that split as sequences of comodules to exact sequences of vector spaces.
Similarly V is said to be relative injective if the functor $\mathrm{Hom}_{\mathcal{YD}_A^A}(-,V)$ *transforms exact sequences of Yetter-Drinfeld modules that split as sequences of modules to exact sequences of vector spaces.*

Relative projective Yetter-Drinfeld modules have the following characterization.

Proposition 4.2. *Let P be a Yetter-Drinfeld module over A. The following assertions are equivalent.*

1. *P is relative projective.*
2. *Any surjective morphism of Yetter-Drinfeld modules $f: M \to P$ that admits a section which is a map of comodules admits a section which is a morphism of Yetter-Drinfeld modules.*
3. *P is a direct summand of a free Yetter-Drinfeld module.*

If A is cosemisimple, these conditions are equivalent to P being a projective object of $\mathcal{YD}_A^A$.

Proof. The proof of (1)⇒(2) is similar to the usual one for modules. Assume (2), and consider the surjective Yetter-Drinfeld module morphism $R(P) \boxtimes A \to P$, $x \otimes a \mapsto x \leftarrow a$. The map $P \to R(P) \boxtimes A$, $x \mapsto x \otimes 1$ is an A-colinear section, so by (2) P is indeed, as a Yetter-Drinfeld module, a direct summand of $R(P) \boxtimes A$.
Assume now that P is free, i.e. $P = V \boxtimes A$ for some comodule V, and let

$$0 \to M \xrightarrow{i} N \xrightarrow{p} Q \to 0$$

be an exact sequence of Yetter-Drinfeld modules that splits as a sequence of comodules. The sequence

$$0 \to \mathrm{Hom}_{\mathcal{YD}_A^A}(P,M) \xrightarrow{i-} \mathrm{Hom}_{\mathcal{YD}_A^A}(P,N) \xrightarrow{p-} \mathrm{Hom}_{\mathcal{YD}_A^A}(P,Q)$$

is exact and we have to show the surjectivity of the map on the right. Let $s: Q \to N$ be a morphism of comodules such that $ps = \mathrm{id}_Q$. Let $\varphi \in \mathrm{Hom}_{\mathcal{YD}_A^A}(V \boxtimes A, Q)$, and let $\varphi_0 : V \to Q$ be defined by $\varphi_0(v) = \varphi(v \otimes 1)$: φ_0 is a map of comodules, and so is $s\varphi_0$. Considering now $\widetilde{s\varphi_0} \in \mathrm{Hom}_{\mathcal{YD}_A^A}(V \boxtimes A, N)$, we have $p\widetilde{s\varphi_0} = \varphi$, which gives the expected surjectivity result. Now if $V \boxtimes A \simeq P \oplus M$ as Yetter-Drinfeld modules, then $\mathrm{Hom}_{\mathcal{YD}_A^A}(V \boxtimes A, -) \simeq \mathrm{Hom}_{\mathcal{YD}_A^A}(P,-) \oplus \mathrm{Hom}_{\mathcal{YD}_A^A}(M,-)$, and the usual argument for projective modules work to conclude that P is relative projective.
It is clear that a projective Yetter-Drinfeld module is relative projective, and if A is cosemisimple, a free Yetter-Drinfeld module is a projective object in $\mathcal{YD}_A^A$ (Proposition 3.3 in [10]), hence a direct summand of a free Yetter-Drinfeld module is projective, and so is a relative projective Yetter-Drinfeld module. $\square$

Similarly, relative injective Yetter-Drinfeld modules are characterized as follows. The proof is similar to the one of the previous result, and is left to the reader.

PROPOSITION 4.3. *Let I be a Yetter-Drinfeld module over A. The following assertions are equivalent.*

1. *I is relative injective.*
2. *Any injective morphism of Yetter-Drinfeld modules $f : I \to M$ that admits a section which is a map of modules admits a section which is a morphism of Yetter-Drinfeld modules.*
3. *P is a direct summand of a co-free Yetter-Drinfeld module.*

4.3 YETTER-DRINFELD MODULES AND HOPF BIMODULES

In this subsection we briefly recall the category equivalence between Yetter-Drinfeld modules and Hopf bimodules [49], and check that the notion of relative projective objects (resp. relative injective objects) for Yetter-Drinfeld modules corresponds to that for Hopf bimodules considered in [53].
First recall that a Hopf bimodule over A is an A-bimodule and A-bicomodule M whose respective left and right coactions $\lambda : M \to A \otimes M$ and $\rho : M \to M \otimes A$ are A-bimodule maps. The category of Hopf bimodules over A, whose morphisms are the bimodule and bicomodule maps, is denoted ${}_A^A\mathcal{M}_A^A$.
If M is Hopf bimodule over A, then ${}^{\mathrm{co}A}M = \{x \in M \mid \lambda(x) = 1 \otimes x\}$ is a right subcomodule of M, and inherits a right A-module structure given by $x \leftarrow a = S(a_{(1)}).x.a_{(2)}$, making it into a Yetter-Drinfeld module over A. This defines a functor

$$\begin{aligned} {}_A^A\mathcal{M}_A^A &\longrightarrow \mathcal{YD}_A^A \\ M &\longmapsto {}^{\mathrm{co}A}M \end{aligned}$$

Conversely, starting from a Yetter-Drinfeld module V, one defines a Hopf bimodule structure on $A \otimes V$ as follows. The bimodule structure is given by

$$a.(b \otimes v).c = abc_{(1)} \otimes (v \leftarrow c_{(2)})$$

and the bicomodule structure is given by the following left and right coactions

$$\begin{aligned} \lambda : A \otimes V &\longrightarrow A \otimes A \otimes V \\ a \otimes v &\longmapsto a_{(1)} \otimes a_{(2)} \otimes v \end{aligned} \qquad \begin{aligned} \rho : A \otimes V &\longrightarrow A \otimes V \otimes A \\ a \otimes v &\longmapsto a_{(1)} \otimes v_{(0)} \otimes a_{(2)}v_{(1)} \end{aligned}$$

If $f : V \longrightarrow W$ is a morphism of Yetter-Drinfeld module, then $\mathrm{id}_A \otimes f : A \otimes V \to A \otimes W$ is a morphism of Hopf bimodules, and hence we get a functor

$$\begin{aligned} \mathcal{YD}_A^A &\longrightarrow {}_A^A\mathcal{M}_A^A \\ V &\longmapsto A \otimes V \end{aligned}$$

The two functors just defined are quasi-inverse equivalences, see [49].

Lemma 4.4. *Relative projective (resp. relative injective) objects in $\mathcal{YD}_A^A$ correspond, via the category equivalence $\mathcal{YD}_A^A \simeq {}_A^A\mathcal{M}_A^A$, to relative projective (resp. relative injective) objects of ${}_A^A\mathcal{M}_A^A$ in the sense of [53].*

Proof. Let M be a Hopf bimodule over A. That M is relatively projective means that the functor $\mathrm{Hom}_{{}_A^A\mathcal{M}_A^A}(M, -)$ transforms exact sequences of Hopf bimodules that split as sequences of bicomodules to exact sequences of vector spaces. The proof of the lemma easily reduces to the following statement.
Let $f : V \to W$ be a surjective morphism of Yetter-Drinfeld modules, inducing a surjective morphism of Hopf bimodules $\mathrm{id}_A \otimes f : A \otimes V \to A \otimes W$. Then there exists an A-comodule section to f if and only if there exists an A-bicomodule section to $\mathrm{id}_A \otimes f$.
Indeed, if $s : W \to V$ is A-colinear with $fs = \mathrm{id}_W$, then $\mathrm{id}_A \otimes s : A \otimes W \to A \otimes V$ is A-bicolinear and is a section to $\mathrm{id}_A \otimes f$. Conversely starting with an A-bicolinear map $T : A \otimes W \to A \otimes V$ with $(\mathrm{id}_A \otimes f)T = \mathrm{id}_{A \otimes W}$, then the map $s : W \to V$ defined by $s(w) = \varepsilon \otimes \mathrm{id}_V(T(1 \otimes w))$ is A-colinear, and satisfies $fs = \mathrm{id}_W$.
In a similar manner, that M is relatively injective means that the functor $\mathrm{Hom}_{{}_A^A\mathcal{M}_A^A}(-, M,)$ transforms exact sequences of Hopf bimodules that split as sequences of bimodules to exact sequences of vector spaces. The proof that this corresponds to the notion of relative injective Yetter-Drinfeld module is left to the reader. □

4.4 Adjoint Hopf subalgebras

We now discuss the way to restrict certain free Yetter-Drinfeld to adjoint Hopf subalgebras.

Proposition 4.5. *Let $B \subset A$ be a Hopf subalgebra. The following assertions are equivalent.*

1. *For any $a \in A$ and $b \in B$, we have*
$$a_{(2)} \otimes S(a_{(1)})ba_{(3)} \in A \otimes B$$
2. *For any B-comodule W, we have $\alpha_{V \boxtimes A}(W \boxtimes A) \subset (W \boxtimes A) \otimes B$ so that $W \boxtimes A$ is an object of $\mathcal{YD}_B^B$.*

Proof. (1) $\Rightarrow$ (2) follows from the definition of $\alpha_{V \boxtimes A}$. Conversely, assuming that (2) holds, take $W = B$ the regular B-comodule. Then for any $a \in A$ and $b \in B$, we have
$$b_{(1)} \otimes a_{(2)} \otimes S(a_{(1)})b_{(2)}a_{(3)} \in A \otimes A \otimes B$$
and hence
$$a_{(2)} \otimes S(a_{(1)})ba_{(3)} = a_{(2)} \otimes S(a_1)\varepsilon(b_{(1)})b_{(2)}a_{(3)} \in A \otimes B$$
Thus (1) holds. □

DEFINITION 4.6. *A Hopf subalgebra $B \subset A$ is said to be adjoint if it satisfies the equivalent conditions of Proposition 4.5.*

Very often adjoint Hopf subalgebras are obtained in the following way. Recall that a Hopf algebra map $f : A \to L$ is said to be cocentral if $f(a_{(1)}) \otimes a_{(2)} = f(a_{(2)}) \otimes a_{(1)}$ for any $a \in A$.

PROPOSITION 4.7. *Let $B \subset A$ be a Hopf subalgebra. Assume that there exists a cocentral and surjective Hopf algebra map $p : A \to L$ such that $B = A^{\mathrm{co}L}$. Then $B \subset A$ is an adjoint Hopf subalgebra. Conversely if $B \subset A$ is an adjoint Hopf subalgebra, if A and B have bijective antipodes and if A is faithfully flat as a right B-module, then there exists a cocentral surjective Hopf algebra map $p : A \to L$ such that $B = A^{\mathrm{co}L}$.*

Proof. Let $a \in A$ and $b \in B$. Since $p(b) = \varepsilon(b)1$, we have, using the cocentrality of p,

$$\begin{aligned}
\mathrm{id}_A \otimes \mathrm{id}_B \otimes p &\left(a_{(2)} \otimes (S(a_{(1)})ba_{(3)})_{(1)} \otimes (S(a_{(1)})ba_{(3)})_{(2)}\right) \\
&= \mathrm{id}_A \otimes \mathrm{id}_B \otimes p \left(a_{(3)} \otimes S(a_{(2)})b_{(1)}a_{(4)} \otimes S(a_{(1)})b_{(2)}a_{(5)}\right) \\
&= a_{(3)} \otimes S(a_{(2)})b_{(1)}a_{(4)} \otimes pS(a_{(1)})p(b_{(2)})p(a_{(5)}) \\
&= a_{(3)} \otimes S(a_{(2)})ba_{(4)} \otimes pS(a_{(1)})p(a_{(5)}) \\
&= a_{(2)} \otimes S(a_{(1)})ba_{(3)} \otimes 1
\end{aligned}$$

Hence $a_{(2)} \otimes S(a_{(1)})ba_{(3)} \in A \otimes A^{\mathrm{co}L} = A \otimes B$: this shows that $B \subset A$ is adjoint.
Conversely, assume that $B \subset A$ is adjoint, that A and B have bijective antipodes and that A is faithfully flat as a right B-module. Then for any $a \in A$ and $b \in B$, we have

$$S(a_{(1)})ba_{(2)} = \varepsilon(a_{(2)})\varepsilon(b_{(1)})S(a_{(1)})b_{(2)}a_{(3)} \in B$$

It is well-known that this implies $B^+A \subset AB^+$, and hence $AB^+ \subset B^+A$ by the bijectivity of the antipodes. It follows that B^+A is a Hopf ideal in A, and we denote by $p : A \to A/B^+A = L$ the canonical Hopf algebra surjection. By construction we have $B \subset A^{\mathrm{co}L}$, and for $b \in B$ we have $p(b) = \varepsilon(b)$. Hence we have for any $a \in A$, $a \otimes 1 = a_{(2)} \otimes p(S(a_{(1)})a_{(3)})$, hence

$$\begin{aligned}
a_{(2)} \otimes p(a_{(1)}) &= (1 \otimes p(a_{(1)}))(a_{(2)} \otimes 1) \\
&= (1 \otimes p(a_{(1)})(a_{(3)} \otimes p(S(a_{(2)})a_{(4)})) = a_{(1)} \otimes p(a_{(2)})
\end{aligned}$$

and this shows that p is cocentral. Finally we have $B = A^{\mathrm{co}L}$ by Corollary 1.8 in [51]. $\square$

We now discuss a condition that ensures that the restriction of a free Yetter-Drinfeld module to an adjoint Hopf subalgebra as in Proposition 4.5 remains a relative projective Yetter-Drinfeld module.

Proposition 4.8. *Let $B \subset A$ be a Hopf subalgebra with $B = A^{\mathrm{co}L}$ for some cocentral and surjective Hopf algebra map $p : A \to L$. Assume that there exists a linear map $\sigma : L \to A$ such that*

1. $p\sigma = \mathrm{id}_L$;
2. $\sigma(x)_{(1)} \otimes p(\sigma(x)_{(2)}) = \sigma(x_{(1)}) \otimes x_{(2)}$, *for any* $x \in L$;
3. $\sigma(x)_{(1)} S(\sigma(x)_{(3)}) \otimes \sigma(x)_{(2)} = 1_B \otimes \sigma(x)$, *for any* $x \in L$.

Then for any B-comodule W, the object $W \boxtimes A \in \mathcal{YD}_B^B$ is relative projective. Such a map σ exists if A is cosemisimple.

Proof. We first claim that for any $a \in A$, we have

$$\sigma p(a_{(1)})_{(1)} \otimes S(\sigma p(a_{(1)})_{(2)}) a_{(2)} \in A \otimes B$$

For any $x \in L$, we have, by (2)

$$\sigma(x)_{(1)} \otimes \sigma(x)_{(2)} \otimes p(\sigma(x)_{(3)}) = \sigma(x_{(1)})_{(1)} \otimes \sigma(x_{(1)})_{(2)} \otimes x_{(2)}$$

and hence for any $a \in A$

$$\sigma p(a)_{(1)} \otimes \sigma p(a)_{(2)} \otimes p(\sigma p(a)_{(3)}) = \sigma p(a_{(1)})_{(1)} \otimes \sigma p(a_{(1)})_{(2)} \otimes p(a_{(2)})$$

We have

$$\begin{aligned}
&(\mathrm{id}_A \otimes p \otimes \mathrm{id}_A)(\mathrm{id}_A \otimes \Delta)\left(\sigma p(a_{(1)})_{(1)} \otimes S(\sigma p(a_{(1)})_{(2)}) a_{(2)}\right)\\
&\quad = \sigma p(a_{(1)})_{(1)} \otimes Sp(\sigma p(a_{(1)})_{(3)}) p(a_{(2)}) \otimes S(\sigma p(a_{(1)})_{(2)}) a_{(3)}\\
&\quad = \sigma p(a_{(1)})_{(1)} \otimes Sp(a_{(2)}) p(a_{(3)}) \otimes S(\sigma p(a_{(1)})_{(2)}) a_{(4)}\\
&\quad = \sigma p(a_{(1)})_{(1)} \otimes 1 \otimes S(\sigma p(a_{(1)})_{(2)}) a_{(2)}
\end{aligned}$$

and this proves our claim.
We thus get for any B-comodule W, a linear map

$$\begin{aligned}
\iota : W \boxtimes A &\longrightarrow (W \boxtimes A) \boxtimes B\\
w \otimes a &\longmapsto w \otimes \sigma p(a_{(1)})_{(1)} \otimes S(\sigma p(a_{(1)})_{(2)}) a_{(2)}
\end{aligned}$$

that we claim to be a morphism of Yetter-Drinfeld modules over B. That ι is a left B-module map is easily checked. Denoting by β the B-coaction on $(W \boxtimes A) \boxtimes B$, we have

$$\begin{aligned}
\beta\iota(w \otimes a) = {} & w_{(0)} \otimes \sigma p(a_{(1)})_{(2)} \otimes S(\sigma p(a_{(1)})_{(5)}) a_{(3)} \otimes\\
& S\left(S(\sigma p(a_{(1)})_{(6)} a_{(2)}\right) S(\sigma p(a_{(1)})_{(1)}) w_{(1)} \sigma p(a_{(1)})_{(3)} S(\sigma p(a_{(1)})_{(4)}) a_{(4)}\\
= {} & w_{(0)} \otimes \sigma p(a_{(1)})_{(2)} \otimes S(\sigma p(a_{(1)})_{(3)}) a_{(3)} \otimes\\
& S\left(S(\sigma p(a_{(1)})_{(4)}) a_{(2)}\right) S(\sigma p(a_{(1)})_{(1)}) w_{(1)} a_{(4)}\\
= {} & w_{(0)} \otimes \sigma p(a_{(1)})_{(2)} \otimes S(\sigma p(a_{(1)})_{(3)}) a_{(3)} \otimes\\
& S(a_{(2)}) S\left(\sigma p(a_{(1)})_1 S(\sigma p(a_{(1)})_{(4)})\right) w_{(1)} a_{(4)}
\end{aligned}$$

By (3), for $x \in L$, we have

$$\sigma(x)_{(2)} \otimes \sigma(x)_{(1)} S(\sigma(x)_{(3)}) = \sigma(x) \otimes 1_B$$

and hence

$$\sigma(x)_{(2)} \otimes S(\sigma(x)_{(3)}) \otimes \sigma(x)_{(1)} S(\sigma(x)_{(4)}) = \sigma(x)_{(1)} \otimes S(\sigma(x)_{(2)}) \otimes 1_B$$

Thus

$$\beta\iota(w \otimes a) = w_{(0)} \otimes \sigma p(a_{(1)})_{(1)} \otimes S(\sigma p(a_{(1)})_{(2)}) a_{(3)} \otimes S(a_{(2)}) w_{(1)} a_{(4)}$$

Now let γ be the B-coaction on $W \boxtimes A$. We have

$$\begin{aligned}(\iota \otimes \mathrm{id}_B)\gamma(w \otimes a) &= \iota \otimes \mathrm{id}_B(w_{(0)} \otimes a_{(2)} \otimes S(a_{(1)}) w_{(1)} a_{(3)}) \\ &= w_{(0)} \otimes \sigma p(a_{(2)})_{(1)} \otimes S(\sigma p(a_{(2)})_{(2)}) a_{(3)} \otimes S(a_{(1)}) w_{(1)} a_{(4)} \\ &= w_{(0)} \otimes \sigma p(a_{(1)})_{(1)} \otimes S(\sigma p(a_{(1)})_{(2)}) a_{(3)} \otimes S(a_{(2)}) w_{(1)} a_{(4)} = \beta\iota(w \otimes a)\end{aligned}$$

where we have used the cocentrality of p. It follows that ι is B-colinear, and hence is a morphism of Yetter-Drinfeld modules over B. Consider now

$$\begin{aligned}\mu : (W \boxtimes A) \boxtimes B &\longrightarrow W \boxtimes A \\ w \otimes a \otimes b &\longmapsto w \otimes ab\end{aligned}$$

It is straightforward to check that μ is a morphism of Yetter-Drinfeld modules over B, with $\mu\iota = \mathrm{id}_{W \boxtimes A}$ and hence we conclude from Proposition 4.2 that $W \boxtimes A$ is a relative projective Yetter-Drinfeld module over B.
For the last assertion, note that L and A both admit right $B^{\mathrm{cop}} \otimes L$-comodule structures given by

$$\begin{aligned}L &\longrightarrow L \otimes (B^{\mathrm{cop}} \otimes L), & A &\longrightarrow A \otimes (B^{\mathrm{cop}} \otimes L) \\ x &\longmapsto x_{(1)} \otimes 1 \otimes x_{(2)}, & a &\longmapsto a_{(2)} \otimes a_{(1)} S(a_{(3)}) \otimes p(a_{(4)})\end{aligned}$$

and that p is $B^{\mathrm{cop}} \otimes L$-colinear. If A is cosemisimple then so is B and so is L (since p is cocentral), hence $B^{\mathrm{cop}} \otimes L$ is cosemisimple. Thus there exists a $B^{\mathrm{cop}} \otimes L$-colinear section to p, which satisfies our 3 conditions. $\square$

There are also situations where the Hopf algebra in the proposition is not cosemisimple and the map σ still exists, see Section 6.

5 Gerstenhaber-Schack cohomology.

5.1 Generalities.

Let A be a Hopf algebra and let V be a Yetter-Drinfeld module over A. The Gerstenhaber-Schack cohomology of A with coefficients in V, that we denote

$H^*_{\mathrm{GS}}(A,V)$, was introduced in [27, 28] by using an explicit bicomplex. In fact Gerstenhaber-Schack used Hopf bimodules instead of Yetter-Drinfeld modules to define their cohomology, but in view of the equivalence between Hopf bimodules and Yetter-Drinfeld modules, we shall work with the simpler framework of Yetter-Drinfeld modules (a Yetter-Drinfeld version of the Gerstenhaber-Schack bicomplex is provided in [45]). A special instance of Gerstenhaber-Schack cohomology is bialgebra cohomology, given by $H^*_b(A) = H^*_{\mathrm{GS}}(A,\mathbb{C})$.
As an example, we have by [46], $H^*_b(\mathbb{C}\Gamma) \simeq H^*(\mathbb{C}\Gamma,\mathbb{C})$ for any discrete group Γ. The bialgebra cohomology of $\mathcal{O}(G)$ for a connected reductive algebraic group G is also described in [46], Theorem 9.2, and some finite-dimensional examples are computed in [58]. Applications to deformations of pointed Hopf algebras are given in [41].
A key result, due to Taillefer [57, 56], shows that Gerstenhaber-Schack cohomology is in fact an Ext-functor:

$$H^*_{\mathrm{GS}}(A,V) \simeq \mathrm{Ext}^*_{\mathcal{YD}^A_A}(\mathbb{C},V)$$

We will use this description as a definition (we will recall and use the definition based on a bicomplex in the proof of the forthcoming Proposition 5.3).

Definition 5.1. *The Gerstenhaber-Schack cohomological dimension of a Hopf algebra A is defined to be*

$$\mathrm{cd}_{\mathrm{GS}}(A) = \sup\{n : \ H^n_{\mathrm{GS}}(A,V) \neq 0 \text{ for some } V \in \mathcal{YD}^A_A\} \in \mathbb{N}\cup\{\infty\}$$

If A and B are Hopf algebras having equivalent tensor categories of comodules, then the given tensor equivalence $F : \mathcal{M}^A \simeq^{\otimes} \mathcal{M}^B$ induces a tensor equivalence $\widehat{F} : \mathcal{YD}^A_A \simeq^{\otimes} \mathcal{YD}^B_B$ (see e.g. [11, 10], this is easy to see thanks to the description of the category of Yetter-Drinfeld modules as the weak center of the category of comodules, see [50]). Hence we get, for any Yetter-Drinfeld module V over A, an isomorphism

$$H^*_{\mathrm{GS}}(A,V) \simeq H^*_{\mathrm{GS}}(B,\widehat{F}(V))$$

and moreover $\mathrm{cd}_{\mathrm{GS}}(A) = \mathrm{cd}_{\mathrm{GS}}(B)$. These properties are what we call the monoidal invariance of Gerstenhaber-Schack cohomology.

5.2 Complexes to compute Gerstenhaber-Schack cohomology.

We now discuss the description of complexes that compute Gerstenhaber-Schack cohomology in particular cases.
Recall that a Hopf algebra A is said to be co-Frobenius if there exists a non-zero A-colinear map $A \to \mathbb{C}$. By [39], A is co-Frobenius if and only if the category $\mathcal{M}^A$ of right comodules has enough projectives. Finite-dimensional Hopf algebras are co-Frobenius, as well as cosemisimple Hopf algebras. See [1, 2] for more examples.

Proposition 5.2. *Let A be a co-Frobenius Hopf algebra and let*

$$\mathbf{P}_{.} = \cdots P_{n+1} \to P_n \to \cdots \to P_1 \to P_0 \to 0$$

be a resolution of $\mathbb{C}$ by projective objects of $\mathcal{YD}_A^A$. We have, for any Yetter-Drinfeld module V over A, an isomorphism

$$H^*_{\mathrm{GS}}(A,V) \simeq H^*(\mathrm{Hom}_{\mathcal{YD}_A^A}(\mathbf{P}_{.},V))$$

and we have

$$\mathrm{cd}_{\mathrm{GS}}(A) = \inf\{n : \ \mathbb{C} \text{ admits a projective resolution of length } n \text{ in } \mathcal{YD}_A^A\}$$

Proof. We know, since A is co-Frobenius, that $\mathcal{YD}_A^A$ has enough projective objects (Corollary 3.4 in [10]). Thus the description of $H^*_{\mathrm{GS}}(A,-)$ as an Ext functor [57] yields that if $\mathbf{P}_{.}$ is a a resolution of $\mathbb{C}$ by projective objects of $\mathcal{YD}_A^A$, we have

$$H^*_{\mathrm{GS}}(A,V) \simeq H^*(\mathrm{Hom}_{\mathcal{YD}_A^A}(\mathbf{P}_{.},V))$$

for any Yetter-Drinfeld module V. The proof of the last statement is similar to the one for group cohomology, see [14, Chapter VIII, Lemma 2.1]. $\square$

Recall [10] that for any $n \in \mathbb{N}$, the Yetter-Drinfeld module $A^{\boxtimes n}$ is defined as follows:

$$A^{\boxtimes 0} = \mathbb{C}, \ A^{\boxtimes 1} = \mathbb{C} \boxtimes A = A_{\mathrm{coad}}, \ A^{\boxtimes 2} = A^{\boxtimes 1} \boxtimes A, \ \ldots, A^{\boxtimes (n+1)} = A^{\boxtimes n} \boxtimes A, \ldots$$

After the obvious vector space identification of $A^{\boxtimes n}$ with $A^{\otimes n}$, the right A-module structure of $A^{\boxtimes n}$ is given by right multiplication and its comodule structure is given by

$$\begin{aligned} \mathrm{ad}_r^{(n)} : A^{\boxtimes n} &\longrightarrow A^{\boxtimes n} \otimes A \\ a_1 \otimes \cdots \otimes a_n &\longmapsto a_{1(2)} \otimes \cdots \otimes a_{n(2)} \otimes S(a_{1(1)} \cdots a_{n(1)}) a_{1(3)} \cdots a_{n(3)} \end{aligned}$$

Proposition 5.3. *Let A be a Hopf algebra and let V be a Yetter-Drinfeld module over A. Assume that one of the following conditions holds.*

1. *A is cosemisimple.*
2. *V is relative injective.*

Then the Gerstenhaber-Schack cohomology $H^*_{\mathrm{GS}}(A,V)$ is the cohomology of the complex

$$\begin{aligned} 0 \to \mathrm{Hom}^A(\mathbb{C},V) \xrightarrow{\partial} &\mathrm{Hom}^A(A^{\boxtimes 1},V) \xrightarrow{\partial} \cdots \\ &\xrightarrow{\partial} \mathrm{Hom}^A(A^{\boxtimes n},V) \xrightarrow{\partial} \mathrm{Hom}^A(A^{\boxtimes n+1},V) \xrightarrow{\partial} \cdots \end{aligned}$$

where the differential $\partial : \mathrm{Hom}^A(A^{\boxtimes n}, V) \longrightarrow \mathrm{Hom}^A(A^{\boxtimes n+1}, V)$ *is given by*

$$\begin{aligned}\partial(f)(a_1 \otimes \cdots \otimes a_{n+1}) =& \varepsilon(a_1) f(a_2 \otimes \cdots \otimes a_{n+1}) + \\ & \sum_{i=1}^{n} (-1)^i f(a_1 \otimes \cdots \otimes a_i a_{i+1} \otimes \cdots \otimes a_{n+1}) \\ & + (-1)^{n+1} f(a_1 \otimes \cdots \otimes a_n) \cdot a_{n+1}\end{aligned}$$

Proof. By [10], Proposition 3.6, the standard resolution of $\mathbb{C}_\varepsilon$ yields in a fact resolution of $\mathbb{C}$ by free Yetter-Drinfeld modules in the category $\mathcal{YD}_A^A$

$$\cdots \longrightarrow A^{\boxtimes n+1} \longrightarrow A^{\boxtimes n} \longrightarrow \cdots \longrightarrow A^{\boxtimes 2} \longrightarrow A^{\boxtimes 1} \longrightarrow 0$$

where each differential is given by

$$\begin{aligned} A^{\boxtimes n+1} &\longrightarrow A^{\boxtimes n} \\ a_1 \otimes \cdots \otimes a_{n+1} &\longmapsto \varepsilon(a_1) a_2 \otimes \cdots \otimes a_{n+1} + \\ & \sum_{i=1}^{n} (-1)^i a_1 \otimes \cdots \otimes a_i a_{i+1} \otimes \cdots \otimes a_{n+1}\end{aligned}$$

If A is cosemimple, then free Yetter-Drinfeld modules are projective, and we get, after standard identification using the fact that the free functor is left adjoint, the result by Proposition 5.2.

To prove the result if the second condition holds, we recall the definition of Gerstenhaber-Schack cohomology using a bicomplex [53]. Let V, W be objects in $\mathcal{YD}_A^A$, let $P_\bullet \to V \to 0$ be a relative projective resolution of V (this means that the objects P_q, $q \geq 0$, are relative projective and the the sequence $P_\bullet \to V \to 0$ splits as a sequence of comodules), and let $0 \to W \to I^\bullet$ be a relative injective resolution of W (this means that the objects I^p, $p \geq 0$, are relative injective and the the sequence $0 \to W \to I^\bullet$ splits as a sequence of modules). We then can form, in a standard way, the bicomplex $C^{\bullet,\bullet}(V, W) = \mathrm{Hom}_{\mathcal{YD}_A^A}(P_\bullet, I^\bullet)$. The uniqueness, up to homotopy, of the previous resolutions ([53], chapter 10) shows that the cohomology of the bicomplex $C^{\bullet,\bullet}(V, W) = \mathrm{Hom}_{\mathcal{YD}_A^A}(P_\bullet, I^\bullet)$ is independent of the choice of these resolutions, and is the Gertenhaber-Schack cohomology of the Yetter-Drinfeld modules V and W (see [56, 57] as well). When $V = \mathbb{C}$, we get the Gerstenhaber Schack-cohomology $H^*_{\mathrm{GS}}(A, W)$ as defined in Subsection 5.1, by [57].

Assuming that W is relative injective, we can use the relative injective resolution

$$0 \to W \to W \to 0 \to \cdots \to 0 \cdots$$

together with the standard resolution of the trivial object $\mathbb{C}$ as above (which is indeed a relative projective resolution of $\mathbb{C}$), and we get a bicomplex with only one non-zero column, which is, again using the fact that the free functor is left adjoint, easily identified with the complex in the statement of the proposition. $\square$

Remark 5.4. When $V = \mathbb{C}$ is the trivial Yetter-Drinfeld module, the complex in Theorem 5.3 is the same as the one defined in [26] in the study of additive deformations of Hopf algebras, which are of interest in quantum probability. This complex is also the complex that defines the so-called Davydov-Yetter cohomology of the tensor category $\mathcal{M}^A$ ([18, 60], see [24], Chapter 7).

Remark 5.5. Let V be a Yetter-Drinfeld module over A. The complex in Proposition 5.3 is a subcomplex of the complex that computes the Hochshild cohomology $H^*(A, {}_\varepsilon V)$, where the left A-module structure on ${}_\varepsilon V$ is the one induced by the counit and the right module structure is the original one. We thus have a linear map

$$H^*_{\mathrm{GS}}(A, V) \to H^*(A, {}_\varepsilon V) \simeq \mathrm{Ext}_A(\mathbb{C}_\varepsilon, V)$$

which is not injective in general. Indeed for $q \in \mathbb{C}^*$ generic ($q = \pm 1$ or not a root of unity), we have $H^3_{\mathrm{GS}}(\mathcal{O}(\mathrm{SL}_q(2)), \mathbb{C}) \simeq \mathbb{C}$ (see [10]), while $H^3(\mathcal{O}(\mathrm{SL}_q(2)), {}_\varepsilon\mathbb{C}_\varepsilon) = 0$ if $q^2 \neq 1$ (see e.g. [31]). In Subsection 5.4 we provide some conditions that ensure that the above map is injective.

5.3 Relation with Hochschild cohomology

We are ready to show that the Gerstenhaber-Schack cohomology of a Hopf algebra determines its Hochschild cohomology.

Theorem 5.6. *Let A be a Hopf algebra and let M be an A-bimodule. Endow $M \otimes A$ with a Yetter-Drinfeld module structure defined by ($a, b \in A$, $m \in M$)*

$$m \otimes a \mapsto m \otimes a_{(1)} \otimes a_{(2)}, \ (m \otimes a) \leftarrow b = S(b_{(2)}).m.b_{(3)} \otimes S(b_{(1)})ab_{(4)}$$

and denote by $M'\#A$ the resulting Yetter-Drinfeld module. Then we have an isomorphism

$$H^*(A, M) \simeq H^*_{\mathrm{GS}}(A, M'\#A)$$

In particular we have $\mathrm{cd}(A) \leq \mathrm{cd}_{\mathrm{GS}}(A)$.

Proof. The Yetter-Drinfeld module $M'\#A$ is the co-free Yetter-Drinfeld associated to the right A-module M' of Section 2. It is thus a relative injective Yetter-Drinfeld module (Proposition 4.3), and we can use the complex of Proposition 5.3 to compute its Gerstenhaber-Schack cohomology.
Recall that since $H^*(A, M) \simeq \mathrm{Ext}^*_A(\mathbb{C}_\varepsilon, M')$ (Section 2), the complex to compute $H^*(A, M)$ is

$$\begin{aligned} 0 \longrightarrow &\mathrm{Hom}(\mathbb{C}, M') \xrightarrow{\partial} \mathrm{Hom}(A, M') \xrightarrow{\partial} \cdots \\ &\cdots \xrightarrow{\partial} \mathrm{Hom}(A^{\otimes n}, M') \xrightarrow{\partial} \mathrm{Hom}(A^{\otimes n+1}, M') \xrightarrow{\partial} \cdots \end{aligned}$$

where the differential $\partial : \mathrm{Hom}(A^{\otimes n}, M') \longrightarrow \mathrm{Hom}(A^{\otimes n+1}, M')$ is given by

$$\begin{aligned}\partial(f)(a_1 \otimes \cdots \otimes a_{n+1}) =& \varepsilon(a_1) f(a_2 \otimes \cdots \otimes a_{n+1}) \\ &+ \sum_{i=1}^{n} (-1)^i f(a_1 \otimes \cdots \otimes a_i a_{i+1} \otimes \cdots \otimes a_{n+1}) \\ &+ (-1)^{n+1} S(a_{n+1(1)}) \cdot f(a_1 \otimes \cdots \otimes a_n) \cdot a_{n+1(2)}\end{aligned}$$

For all $n \geq 0$, we have linear isomorphisms

$$\begin{aligned}\mathrm{Hom}^A(A^{\boxtimes n}, M'\#A) &\longrightarrow \mathrm{Hom}(A^{\otimes n}, M') \\ f &\longmapsto (\mathrm{id}_M \otimes \varepsilon) f\end{aligned}$$

For $f \in \mathrm{Hom}(A^{\boxtimes n}, M'\#A)$ and $a_1, \ldots, a_n \in A$, with $f(a_1 \otimes \cdots \otimes a_n) = \sum_i m_i \otimes b_i$, we have

$$\begin{aligned}\mathrm{id}_M &\otimes \varepsilon (f(a_1 \otimes \cdots \otimes a_n) \leftarrow a_{n+1}) \\ &= \mathrm{id}_M \otimes \varepsilon \left(\sum_i S(a_{n+1(2)}).m_i.a_{n+1(3)} \otimes S(a_{n+1(1)}) b_i a_{n+1(4)} \right) \\ &= \sum_i \varepsilon(b_i) S(a_{n+1(1)}).m_i.a_{n+1(2)} \\ &= S(a_{n+1(1)}).\left((\mathrm{id}_M \otimes \varepsilon)(f(a_1 \otimes \cdots \otimes a_{n+1})\right).a_{n+1(2)}\end{aligned}$$

From this computation it follows easily that the previous isomorphisms commute with the differentials (as already said, the one for Gerstenhaber-Schack cohomology being given by the complex of Proposition 5.3), and hence the complexes that define both cohomologies are isomorphic. ∎

We get the results announced in the introduction, providing a partial answer to Question 1.1.

Corollary 5.7. *Let A and B be Hopf algebras such that there exists an equivalence of linear tensor categories $\mathcal{M}^A \simeq^{\otimes} \mathcal{M}^B$. Then there exist two functors*

$$F_1 : {}_A\mathcal{M}_A \to \mathcal{YD}_B^B \text{ and } F_2 : {}_B\mathcal{M}_B \to \mathcal{YD}_A^A$$

such that for any A-bimodule M and any B-bimodule N, we have

$$H^*(A, M) \simeq H^*_{\mathrm{GS}}(B, F_1(M)) \text{ and } H^*(B, N) \simeq H^*_{\mathrm{GS}}(A, F_2(N))$$

In particular we have $\max(\mathrm{cd}(A), \mathrm{cd}(B)) \leq \mathrm{cd}_{\mathrm{GS}}(A) = \mathrm{cd}_{\mathrm{GS}}(B)$.

Proof. The construction in the previous theorem clearly yields a functor ${}_A\mathcal{M}_A \to \mathcal{YD}_A^A$, that we compose with the functor $\mathcal{YD}_A^A \to \mathcal{YD}_B^B$ from the discussion at the end of subsection 5.1, to get the announced functor F_1, and similarly the functor F_2. The last claim follows immediately. ∎

Remark 5.8. Recall that Question 1.2, motivated by Theorem 5.6, asks if $\mathrm{cd}(A) = \mathrm{cd}_{\mathrm{GS}}(A)$ for any Hopf algebra A. Question 1.2 has indeed a positive answer in the finite-dimensional case: if A is semisimple, then it is cosemisimple by the Larson-Radford theorem [38], and hence $\mathcal{YD}_A^A$ is semisimple (since the Drinfeld double $D(A)$ is then semisimple, see [48]), so we have $\mathrm{cd}(A) = 0 = \mathrm{cd}_{\mathrm{GS}}(A)$. If A is not semisimple, then $\mathrm{cd}(A) = \infty = \mathrm{cd}_{\mathrm{GS}}(A)$. It thus follows that a positive answer to Question 1.2 would provide a natural infinite-dimensional generalization to the above mentioned Larson-Radford theorem.

The characteristic zero assumption is indeed necessary: if A is a finite-dimensional semisimple non cosemisimple Hopf algebra, the base field being then necessarily of characteristic > 0 [38], then $\mathrm{cd}(A) = 0 < \mathrm{cd}_{\mathrm{GS}}(A) = \infty$.

See the next subsection for some partial results in the cosemisimple case.

5.4 Cosemisimple Hopf algebras

We now provide some more precise partial answers to Questions 1.1 and 1.2 when the Hopf algebra is cosemisimple and of Kac type (recall that this means that $S^2 = \mathrm{id}$).

Proposition 5.9. *Let A be a cosemisimple Hopf algebra of Kac type, and let V be a Yetter-Drinfeld module over A. Then the natural linear map*

$$H^*_{\mathrm{GS}}(A, V) \to H^*(A, {}_\varepsilon V)$$

arising from Proposition 5.3 is injective.

Proof. Let h be the Haar integral on A. Recall that for any A-comodules V and W, we have a surjective averaging operator

$$\begin{aligned} M : \mathrm{Hom}(V, W) &\longrightarrow \mathrm{Hom}^A(V, W) \\ f &\longmapsto M(f),\ M(f)(v) = h\left(f(v_{(0)})_{(1)} S(v_{(1)})\right) f(v_{(0)})_{(0)} \end{aligned}$$

with $f \in \mathrm{Hom}^A(V, W)$ if and only if $M(f) = f$. Now let V be our given Yetter-Drinfeld module, and let $f \in \mathrm{Hom}(A^{\otimes n}, V)$. We thus have $M(f) \in \mathrm{Hom}^A(A^{\boxtimes n}, V)$, with

$$\begin{aligned} &M(f)(a_1 \otimes \cdots \otimes a_n) = \\ &\qquad h\left(f(a_{1(2)} \otimes \cdots \otimes a_{n(2)})_{(1)} S(a_{1(3)} \cdots a_{n(3)}) S^2(a_{1(1)} \cdots a_{n(1)})\right) \\ &\qquad f(a_{1(2)} \otimes \cdots \otimes a_{n(2)})_{(0)} \end{aligned}$$

It is a tedious but straightforward verification to check that, under our assumption, we have $\partial(M(f)) = M(\partial(f))$. To convince the reader, we present the verification at $n=2$. Let $f \in \mathrm{Hom}(A^{\otimes 2}, V)$. We have

$$\begin{aligned}
&\partial(M(f))(a\otimes b\otimes c) = \\
&\quad \varepsilon(a)h\left(f(b_{(2)}\otimes c_{(2)})_{(1)}S(b_{(3)}c_{(3)})S^2(b_{(1)}c_{(1)})\right)f(b_{(2)}\otimes c_{(2)})_{(0)} \\
&\quad - h\left(f(a_{(2)}b_{(2)}\otimes c_{(2)})_{(1)}S(a_{(3)}b_{(3)}c_{(3)})S^2(a_{(1)}b_{(1)}c_{(1)})\right)f(a_{(2)}b_{(2)}\otimes c_{(2)})_{(0)} \\
&\quad + h\left(f(a_{(2)}\otimes b_{(2)}c_{(2)})_{(1)}S(a_{(3)}b_{(3)}c_{(3)})S^2(a_{(1)}b_{(1)}c_{(1)})\right)f(a_{(2)}\otimes b_{(2)}c_{(2)})_{(0)} \\
&\quad - h\left(f(a_{(2)}\otimes b_{(2)})_{(1)}S(a_{(3)}b_{(3)})S^2(a_{(1)}b_{(1)})\right)f(a_{(2)}\otimes b_{(2)})_{(0)}\cdot c
\end{aligned}$$

On the other hand we have

$$\begin{aligned}
&M(\partial(f))(a\otimes b\otimes c) = \\
&\quad h\left(\partial(f)(a_{(2)}\otimes b_{(2)}\otimes c_{(2)})_{(1)}S(a_{(3)}b_{(3)}c_{(3)})S^2(a_{(1)}b_{(1)}c_{(1)})\right) \\
&\qquad\qquad \partial(f)(a_{(2)}\otimes b_{(2)}\otimes c_{(2)})_{(0)} \\
&\quad = h\left(\varepsilon(a_{(2)})f(b_{(2)}\otimes c_{(2)})_{(1)}S(a_{(3)}b_{(3)}c_{(3)})S^2(a_{(1)}b_{(1)}c_{(1)})\right)f(b_{(2)}\otimes c_{(2)})_{(0)} \\
&\quad - h\left(f(a_{(2)}b_{(2)}\otimes c_{(2)})_{(1)}S(a_{(3)}b_{(3)}c_{(3)})S^2(a_{(1)}b_{(1)}c_{(1)})\right)f(a_{(2)}b_{(2)}\otimes c_{(2)})_{(0)} \\
&\quad + h\left(f(a_{(2)}\otimes b_{(2)}c_{(2)})_{(1)}S(a_{(3)}b_{(3)}c_{(3)})S^2(a_{(1)}b_{(1)}c_{(1)})\right)f(a_{(2)}\otimes b_{(2)}c_{(2)})_{(0)} \\
&\quad - h\left((f(a_{(2)}\otimes b_{(2)})\cdot c_{(2)})_{(1)}S(a_{(3)}b_{(3)}c_{(3)})S^2(a_{(1)}b_{(1)}c_{(1)})\right) \\
&\quad ((f(a_{(2)}\otimes b_{(2)})\cdot c_{(2)})_{(0)}
\end{aligned}$$

Using the Yetter-Drinfeld condition, the last expression equals

$$h\left(S(c_{(2)})f(a_{(2)}\otimes b_{(2)})_{(1)}c_{(4)}S(a_{(3)}b_{(3)}c_{(5)})S^2(a_{(1)}b_{(1)}c_{(1)})\right)(f(a_{(2)}\otimes b_{(2)})_{(0)}\cdot c_{(3)}$$

The fact that $S^2=\mathrm{id}$ and that the Haar integral is a trace (since $S^2=\mathrm{id}$) then shows that this last expression equals the last one in the computation of $\partial(M(f))(a\otimes b\otimes c)$, and shows that indeed $\partial(M(f))=M(\partial(f))$.
Now let $f \in \mathrm{Hom}^A(A^{\boxtimes n}, V)$ be such that $f=\partial(\mu)$ for some $\mu \in \mathrm{Hom}(A^{\otimes n-1}, V)$. Then $M(f)=M(\partial(\mu))=\partial(M(\mu)))$, with $M(\mu) \in \mathrm{Hom}^A(A^{\boxtimes n-1}, V)$, and hence $f=0$ in $H^n_{\mathrm{GS}}(A,V)$: our claim is proved. ☐

We thus get the following partial answers to Questions 1.2 and 1.1.

Corollary 5.10. *Let A be cosemisimple Hopf algebra of Kac type. Then* $\mathrm{cd}(A)=\mathrm{cd}_{\mathrm{GS}}(A)$.

Proof. We have $\mathrm{cd}(A)\leq \mathrm{cd}_{\mathrm{GS}}(A)$ by Theorem 5.6, and the previous proposition ensures that $\mathrm{cd}_{\mathrm{GS}}(A)\leq \mathrm{cd}(A)$. ☐

Corollary 5.11. *Let A and B be cosemisimple Hopf algebras such that there exists an equivalence of linear tensor categories $\mathcal{M}^A \simeq^{\otimes} \mathcal{M}^B$. If A is of Kac type, then we have $\mathrm{cd}(A)\geq \mathrm{cd}(B)$, and if A and B both are of Kac type, then* $\mathrm{cd}(A)=\mathrm{cd}(B)$.

Proof. We have, combining Theorem 5.6 and the previous corollary,

$$\mathrm{cd}(A) = \mathrm{cd}_{\mathrm{GS}}(A) = \mathrm{cd}_{\mathrm{GS}}(B) \geq \mathrm{cd}(B)$$

with $\mathrm{cd}(B) = \mathrm{cd}_{\mathrm{GS}}(B)$ if B is of Kac type as well. □

See the next section for examples that are not of Kac type.

6 Application to quantum symmetry algebras

In this section we provide applications of the previous considerations to quantum symmetry algebras.

6.1 The universal Hopf algebra of a non-degenerate bilinear form and its adjoint subalgebra

Let $E \in \mathrm{GL}_n(\mathbb{C})$. Recall that the algebra $\mathcal{B}(E)$ [21] is presented by generators $(u_{ij})_{1\leq i,j\leq n}$ and relations

$$E^{-1}u^tEu = I_n = uE^{-1}u^tE,$$

where u is the matrix $(u_{ij})_{1\leq i,j\leq n}$. It has a Hopf algebra structure defined by

$$\Delta(u_{ij}) = \sum_{k=1}^{n} u_{ik} \otimes u_{kj},\ \varepsilon(u_{ij}) = \delta_{ij},\ S(u) = E^{-1}u^tE$$

The Hopf algebra $\mathcal{B}(E)$ represents the quantum symmetry group of the bilinear form associated to the matrix E. It can also be constructed as a quotient of the FRT bialgebra associated to Yang-Baxter operators constructed by Gurevich [30]. For the matrix

$$E_q = \begin{pmatrix} 0 & 1 \\ -q^{-1} & 0 \end{pmatrix}$$

we have $\mathcal{B}(E_q) = \mathcal{O}(\mathrm{SL}_q(2))$, and thus the Hopf algebras $\mathcal{B}(E)$ are natural generalizations of $\mathcal{O}(\mathrm{SL}_q(2))$. It is shown in [9] that for $q \in \mathbb{C}^*$ satisfying $\mathrm{tr}(E^{-1}E^t) = -q - q^{-1}$, the tensor categories of comodules over $\mathcal{B}(E)$ and $\mathcal{O}(\mathrm{SL}_q(2))$ are equivalent. Thus $\mathcal{B}(E)$ is cosemisimple if and only if the corresponding q is not a root of unity or $q = \pm 1$.
It was proved in [10] that if $n \geq 2$, then $\mathrm{cd}(\mathcal{B}(E)) = 3$ (Theorem 6.1 and Proposition 6.4 in [10], see e.g. [31] for the case $E = E_q$ and [17] for the case $E = I_n$), and the bialgebra cohomology of $\mathcal{B}(E)$ was computed there in the cosemisimple case.
As a preliminary step towards the study of quantum symmetry algebras of semisimple algebras, we now study the adjoint subalgebra $\mathcal{B}_+(E)$ of $\mathcal{B}(E)$.
The algebra $\mathcal{B}_+(E)$ is, by definition, the subalgebra of $\mathcal{B}(E)$ generated by the elements $u_{ij}u_{kl}$, $1 \leq i,j,k,l \leq n$. It is easily seen to be a Hopf subalgebra. Also it is easily seen that $\mathcal{B}_+(E) = \mathcal{B}(E)^{\mathrm{co}\mathbb{C}\mathbb{Z}_2}$, where p is the cocentral Hopf

algebra map $\mathcal{B}(E) \to \mathbb{C}\mathbb{Z}_2$, $u_{ij} \mapsto \delta_{ij} g$, where g stands for the generator of $\mathbb{Z}_2$, the cyclic group of order 2. The Hopf algebra $\mathcal{B}_+(E)$ is cosemisimple if and only if $\mathcal{B}(E)$ is.

Lemma 6.1. *Assume that* $\mathrm{tr}(E^{-1}E^t) \neq 0$. *Then there exists a linear map* $\sigma : \mathbb{C}\mathbb{Z}_2 \to \mathcal{B}(E)$ *satisfying the conditions of Proposition 4.8.*

Proof. Consider the matrix $F = E(E^t)^{-1} = (\alpha_{ij})$. We have $\mathrm{tr}(F) = \mathrm{tr}(E^{-1}E^t) = t \neq 0$. Consider the element $x = t^{-1}\sum_{ij} \alpha_{ij} u_{ij} \in \mathcal{B}(E)$ and let $\sigma : \mathbb{C}\mathbb{Z}_2 \to \mathcal{B}(E)$ be the unique linear map such that $\sigma(1) = 1$ and $\sigma(g) = x$. It is straightforward to check that σ indeed satisfies the conditions of Proposition 4.8. $\square$

Theorem 6.2. *Let* $E \in \mathrm{GL}_n(\mathbb{C})$ *with* $n \geq 2$. *Then we have* $\mathrm{cd}(\mathcal{B}_+(E)) = 3 \leq \mathrm{cd}_{\mathrm{GS}}(\mathcal{B}_+(E))$, *and if moreover* $\mathcal{B}_+(E)$ *is cosemisimple, then* $\mathrm{cd}_{\mathrm{GS}}(\mathcal{B}_+(E)) = 3$.

Proof. We have, by Proposition 2.4, a strict exact sequence of Hopf algebras

$$\mathbb{C} \to \mathcal{B}_+(E) \to \mathcal{B}(E) \to \mathbb{C}\mathbb{Z}_2 \to \mathbb{C}$$

so it follows from Proposition 3.2 that $\mathrm{cd}(\mathcal{B}_+(E)) = \mathrm{cd}(\mathcal{B}(E)) = 3$. By Theorem 5.6 we have $\mathrm{cd}_{\mathrm{GS}}(\mathcal{B}_+(E)) \geq 3$.
Consider now the exact sequence of free Yetter-Drinfeld modules over $\mathcal{B}(E)$ from [10]:

$$0 \to \mathbb{C}\boxtimes\mathcal{B}(E) \xrightarrow{\phi_1} (V_E^*\otimes V_E)\boxtimes\mathcal{B}(E) \xrightarrow{\phi_2} (V_E^*\otimes V_E)\boxtimes\mathcal{B}(E) \xrightarrow{\phi_3} \mathbb{C}\boxtimes\mathcal{B}(E) \xrightarrow{\varepsilon} \mathbb{C} \to 0$$

All the $\mathcal{B}(E)$-comodules involved in the left terms are in fact comodules over $\mathcal{B}_+(E)$, so we have, by Proposition 4.5, an exact sequence of Yetter-Drinfeld modules over $\mathcal{B}_+(E)$. Assume now that $\mathcal{B}_+(E)$ is cosemisimple. The previous lemma ensures that we are in the situation of Proposition 4.8, so all the terms in the sequence (except the last one of course) are projective Yetter-Drinfeld modules over $\mathcal{B}_+(E)$. We conclude from Proposition 5.2 that $\mathrm{cd}_{\mathrm{GS}}(\mathcal{B}_+(E)) \leq 3$, and hence that $\mathrm{cd}_{\mathrm{GS}}(\mathcal{B}_+(E)) = 3$. $\square$

To compute the bialgebra cohomology of $\mathcal{B}_+(E)$ in the cosemisimple case, we need some preliminaries. We specialize at $E_q = \begin{pmatrix} 0 & 1 \\ -q^{-1} & 0 \end{pmatrix}$ and we put $A = \mathcal{B}(E_q) = \mathcal{O}(\mathrm{SL}_q(2))$ (with its standard generators a, b, c, d) and $B = \mathcal{B}_+(E_q)$. In the next lemma we only assume that $q + q^{-1} \neq 0$. Recall from Subsection 4.4 that if W is a B-comodule, then $W \boxtimes A$ is a Yetter-Drinfeld module over B.

Lemma 6.3. *We have, for any B-comodule W, a vector space isomorphism*

$$\begin{aligned} \mathrm{Hom}_{\mathcal{YD}_B^B}(W \boxtimes A, \mathbb{C}) &\longrightarrow \mathrm{Hom}^B(W, \mathbb{C}) \oplus \mathrm{Hom}^B(W, \mathbb{C}) \\ \psi &\longmapsto (\psi(- \otimes 1), \psi(- \otimes \chi)) \end{aligned}$$

where $\chi = q^{-1}a + qd$.

Proof. Let $\psi \in \mathrm{Hom}_{\mathcal{YD}_B^B}(W \boxtimes A, \mathbb{C})$. That both $\psi(- \otimes 1)$ and $\psi(- \otimes \chi)$ are B-comodule maps follow from the fact that 1 and χ are coinvariant for the co-adjoint action of A. We have, for any $w \in W$, using the B-linearity

$$\psi(w \otimes b) = \psi(w \otimes b(ad - q^{-1}bc)) = \psi(w \otimes bad) = q\psi(w \otimes abd) = 0$$

and similarly $\psi(w \otimes c) = 0$. We also have

$$\psi(w \otimes d) = \psi(w \otimes d(ad - q^{-1}bc))) = \psi(w \otimes dad) = \psi(w \otimes ad^2) = \psi(w \otimes a)$$

These identities, together with the fact that $A = B \oplus B'$, where $B' = XB$ and $X = \{a, b, c, d\}$, show that the map in the statement of the lemma is injective. For $(\psi_1, \psi_2) \in \mathrm{Hom}^B(W, \mathbb{C}) \oplus \mathrm{Hom}^B(W, \mathbb{C})$, we define a linear map $\psi : W \otimes A \to \mathbb{C}$ by

$$\psi(w \otimes (y + y')) = \psi_1(w)\varepsilon(y) + (q + q^{-1})^{-1}\psi_2(w)\varepsilon(y'), \ y \in B, \ y \in B'$$

It is clear that ψ is A-linear and a direct verification to check that ψ is a map of B-comodules, for the co-action of $W \boxtimes A$. Hence we have $\psi \in \mathrm{Hom}_{\mathcal{YD}_B^B}(W \boxtimes A, \mathbb{C})$, and clearly $\psi(- \otimes 1) = \psi_1$ and $\psi(- \otimes \chi) = \psi_2$. Therefore our map is surjective, and we are done. $\square$

Theorem 6.4. *Let $E \in \mathrm{GL}_n(\mathbb{C})$ with $n \geq 2$. If $\mathcal{B}_+(E)$ is cosemisimple, then*

$$H_b^n(\mathcal{B}_+(E)) \simeq \begin{cases} 0 & \text{if } n \neq 0, 3 \\ \mathbb{C} & \text{if } n = 0, 3 \end{cases}$$

Proof. The monoidal invariance of bialgebra cohomology enables us to assume that $E = E_q$ as in the previous discussion, of which we keep the notations. We denote by V the fundamental A-comodule of dimension 2, of which we fix a basis e_1, e_2. We have an exact sequence of Yetter-Drinfeld modules over A (and over B)

$$0 \to \mathbb{C} \boxtimes A \xrightarrow{\phi_1} (V^* \otimes V) \boxtimes A \xrightarrow{\phi_2} (V^* \otimes V) \boxtimes A \xrightarrow{\phi_3} \mathbb{C} \boxtimes A \xrightarrow{\varepsilon} \mathbb{C} \to 0$$

with for any $x \in A$ (see the proof of Lemma 5.6 in [10])

$$\begin{aligned}
\phi_1(x) = e_1^* \otimes e_1 \otimes &((-q^{-1} + qd)x) + e_1^* \otimes e_2 \otimes (-cx) \\
&+ e_2^* \otimes e_1 \otimes (-bx) + e_2^* \otimes e_2 \otimes ((-q + q^{-1}a)x) \\
\phi_2(e_1^* \otimes e_1 \otimes x) &= e_1^* \otimes e_1 \otimes x + e_2^* \otimes e_1 \otimes (-qbx) + e_2^* \otimes e_2 \otimes ax \\
\phi_2(e_1^* \otimes e_2 \otimes x) &= e_1^* \otimes e_1 \otimes bx + e_1^* \otimes e_2 \otimes (1 - q^{-1}a)x \\
\phi_2(e_2^* \otimes e_1 \otimes x) &= e_2^* \otimes e_1 \otimes (1 - qd)x + e_2^* \otimes e_2 \otimes cx \\
\phi_2(e_2^* \otimes e_2 \otimes x) &= e_1^* \otimes e_1 \otimes dx + e_1^* \otimes e_2 \otimes (-q^{-1}cx) + e_2^* \otimes e_2 \otimes x \\
\phi_3(e_1^* \otimes e_1 \otimes x) &= (a - 1)x, \quad \phi_3(e_1^* \otimes e_2 \otimes x) = bx, \\
\phi_3(e_2^* \otimes e_1 \otimes x) &= cx, \quad \phi_3(e_2^* \otimes e_2 \otimes x) = (d - 1)x
\end{aligned}$$

and by Lemma 6.3, Proposition 4.8 and Proposition 5.2, the bialgebra cohomology of B is the cohomology of the complex

$$\begin{aligned}0 \to &\mathrm{Hom}_{\mathcal{YD}_B^B}(\mathbb{C} \boxtimes A, \mathbb{C}) \xrightarrow{\phi_3^t} \mathrm{Hom}_{\mathcal{YD}_B^B}(V^* \otimes V) \boxtimes A, \mathbb{C}) \\ &\xrightarrow{\phi_2^t} \mathrm{Hom}_{\mathcal{YD}_B^B}(V^* \otimes V) \boxtimes A, \mathbb{C}) \xrightarrow{\phi_1^t} \mathrm{Hom}_{\mathcal{YD}_B^B}(\mathbb{C} \boxtimes A, \mathbb{C}) \to 0\end{aligned}$$

We have, by the previous lemma, $\mathrm{Hom}_{\mathcal{YD}_B^B}(\mathbb{C} \boxtimes A, \mathbb{C}) \simeq \mathbb{C}^2$, and

$$\mathrm{Hom}_{\mathcal{YD}_B^B}(V^* \otimes V) \boxtimes A, \mathbb{C}) \simeq \mathrm{Hom}^B(V^* \otimes V, \mathbb{C}) \oplus \mathrm{Hom}^B(V^* \otimes V, \mathbb{C}) \simeq \mathbb{C}^2$$

Therefore the previous complex is isomorphic to a complex of the form

$$0 \longrightarrow \mathbb{C}^2 \longrightarrow \mathbb{C}^2 \longrightarrow \mathbb{C}^2 \longrightarrow \mathbb{C}^2 \longrightarrow 0$$

The reader will easily write down explicitly this complex and compute its cohomology, yielding the announced result for the bialgebra cohomology of B. $\square$

6.2 Bialgebra cohomology and cohomological dimensions of $A_{\mathrm{aut}}(R, \varphi)$

Let (R, φ) be a finite-dimensional measured algebra: this means that R is a finite-dimensional algebra and $\varphi : R \to \mathbb{C}$ is a linear map (a measure on R) such that the associated bilinear map $R \times R \to \mathbb{C}$, $(x, y) \mapsto \varphi(xy)$ is non-degenerate. Thus a finite-dimensional measured algebra is a Frobenius algebra together with a fixed measure. A coaction of a Hopf algebra A on a finite-dimensional measured algebra (R, φ) is an A-comodule structure on R making it into an A-comodule algebra and such that $\varphi : R \to \mathbb{C}$ is A-colinear. It is well-known that there exists a universal Hopf algebra coacting on (R, φ) (see [61] in the compact case with R semisimple and [8] in general), that we denote $A_{\mathrm{aut}}(R, \varphi)$ and call the quantum symmetry algebra of (R, φ). The following particular cases are of special interest.

1. For $R = \mathbb{C}^n$ and $\varphi = \varphi_n$ the canonical integration map (with $\varphi_n(e_i) = 1$ for $e_1, \dots, e_n$ the canonical basis of $\mathbb{C}^n$), we have $A_{\mathrm{aut}}(\mathbb{C}^n, \varphi_n) =: A_s(n)$, the coordinate algebra on the quantum permutation group [61], presented by generators x_{ij}, $1 \leq i, j \leq n$, submitted to the relations

$$\sum_{l=1}^n x_{li} = 1 = \sum_{l=1}^n x_{il},\ x_{ik}x_{ij} = \delta_{kj}x_{ij},\ x_{ki}x_{ji} = \delta_{kj}x_{ji},\ 1 \leq i, j, k \leq n$$

Its Hopf algebra structure is defined by

$$\Delta(x_{ij}) = \sum_{k=1}^n x_{ik} \otimes x_{kj},\ \varepsilon(x_{ij}) = \delta_{ij},\ S(x_{ij}) = x_{ji}$$

The Hopf algebra $A_s(n)$ is infinite-dimensional if $n \geq 4$ [61].

2. For $R = M_2(\mathbb{C})$ and $q \in \mathbb{C}^*$, let $\mathrm{tr}_q : M_2(\mathbb{C}) \to \mathbb{C}$ be the q-trace, i.e. $\mathrm{tr}_q(g) = qg_{11} + q^{-1}g_{22}$ for $g = (g_{ij}) \in M_2(\mathbb{C})$. Then we have $A_{\mathrm{aut}}(M_2(\mathbb{C}), \mathrm{tr}_q) \simeq \mathcal{O}(\mathrm{PSL}_q(2))$, the latter algebra being $\mathcal{B}_+(E_q)$ in the notation of the previous subsection (it is often denoted $\mathcal{O}(\mathrm{SO}_{q^{1/2}}(3))$, see e.g. [35]). The above isomorphism $A_{\mathrm{aut}}(M_2(\mathbb{C}), \mathrm{tr}_q) \to \mathcal{O}(\mathrm{PSL}_q(2))$ is constructed using the universal property of $A_{\mathrm{aut}}(M_2(\mathbb{C}), \mathrm{tr}_q)$, and the verification that it is indeed injective is a long and tedious computation, as in [20].

Let (R, φ) be a finite-dimensional measured algebra. Since $\varphi \circ m$ is non-degenerate, where m is the multiplication of R, there exists a linear map $\delta : \mathbb{C} \to R \otimes R$ such that $(R, \varphi \circ m, \delta)$ is a left dual for R, i.e.

$$((\varphi \circ m) \otimes \mathrm{id}_R) \circ (\mathrm{id}_R \otimes \delta) = \mathrm{id}_R = (\mathrm{id}_R \otimes (\varphi \circ m)) \circ (\delta \otimes \mathrm{id}_R)$$

Following [43], we put

$$\tilde{\varphi} = \varphi \circ m \circ (m \otimes \mathrm{id}_R) \circ (\mathrm{id}_R \otimes \delta) : R \to \mathbb{C}$$

and we say that (R, φ) (or φ) is normalizable if $\varphi(1) \neq 0$ and if there exists $\lambda \in \mathbb{C}^*$ such that $\tilde{\varphi} = \lambda\varphi$. Using the definition of Frobenius algebra in terms of coalgebras, the coproduct is $\Delta = (m \otimes \mathrm{id}_R) \circ (\mathrm{id}_R \otimes \delta) = (\mathrm{id}_R \otimes m) \circ (\delta \otimes \mathrm{id}_R)$, and we have $\tilde{\varphi} = \varphi \circ m \circ \Delta$.
The condition that φ is normalizable is equivalent to require, in the language of [33, Definition 3.1], that $R/\mathbb{C}$ is a strongly separable extension with Frobenius system (φ, x_i, y_i), where $\delta(1) = \sum_i x_i \otimes y_i$. It thus follows that if φ is normalizable, then R is necesarily a separable (semisimple) algebra. Conversely, if R is semisimple, writing R as a direct product of matrix algebras, one easily sees the conditions that ensure that φ is normalizable, see [43].
It is shown in [43] (Corollary 4.9), generalizing earlier results from [5, 6, 19], that if (R, φ) is a finite-dimensional semisimple measured algebra with $\dim(R) \geq 4$ and φ normalizable, then there exists $q \in \mathbb{C}^*$ with $q + q^{-1} \neq 0$ such that

$$\mathcal{M}^{A_{\mathrm{aut}}(R,\varphi)} \simeq^{\otimes} \mathcal{M}^{\mathcal{O}(\mathrm{PSL}_q(2))}$$

The parameter q is determined as follows. First consider $\lambda \in \mathbb{C}^*$ such that $\tilde{\varphi} = \lambda\varphi$ and choose $\mu \in \mathbb{C}^*$ such that $\mu^2 = \lambda\varphi(1)$. Then q is any solution of the equation $q + q^{-1} = \mu$ (recall that $\mathcal{O}(\mathrm{PSL}_q(2)) = \mathcal{O}(\mathrm{PSL}_{-q}(2))$, so the choice of μ does not play any role).
As an example, for $(\mathbb{C}^n, \varphi_n)$ as above (and $n \geq 4$), φ_n is normalizable with the corresponding λ equal to 1, and q is any solution of the equation $q + q^{-1} = \sqrt{n}$.

Theorem 6.5. *Let (R, φ) be a finite-dimensional semisimple measured algebra with $\dim(R) \geq 4$ and φ normalizable. Assume that $A_{\mathrm{aut}}(R, \varphi)$ is cosemisimple. Then we have*

$$H_b^n(A_{\mathrm{aut}}(R, \varphi)) \simeq \begin{cases} 0 & \text{if } n \neq 0, 3 \\ \mathbb{C} & \text{if } n = 0, 3 \end{cases}$$

and $\mathrm{cd}(A_{\mathrm{aut}}(R,\varphi)) \leq \mathrm{cd}_{\mathrm{GS}}(A_{\mathrm{aut}}(R,\varphi)) = 3$*, with equality if* φ *is a trace. In particular we have* $\mathrm{cd}(A_s(n)) = 3 = \mathrm{cd}_{\mathrm{GS}}(A_s(n))$ *for any* $n \geq 4$*.*

Proof. The proof follows immediately from the combination of the above monoidal equivalence, the monoidal invariance of Gerstenhaber-Schack cohomology, Theorem 6.2, Theorem 6.4, Theorem 5.6, and Corollary 5.10 ($A_{\mathrm{aut}}(R,\varphi)$ being of Kac type when φ is a trace). $\square$

Note that the length 3 resolution of the trivial Yetter-Drinfeld module over $\mathcal{O}(\mathrm{PSL}_q(2))$ by relative projective Yetter-Drinfeld modules considered in the previous subsection (see the proof of Theorem 6.4) transports to a length 3 resolution of the trivial Yetter-Drinfeld module over $A_{\mathrm{aut}}(R,\varphi)$ by relative projective Yetter-Drinfeld modules (see Theorem 4.1 in [10]), and in particular this yields a length 3 projective resolution of the trivial module over $A_{\mathrm{aut}}(R,\varphi)$. We have not been able to write down this resolution explicitly enough to compute Hochschild cohomology groups and show that one always has $\mathrm{cd}(A_{\mathrm{aut}}(R,\varphi)) = 3$. We believe that this is true however.

Remark 6.6. It follows that the L^2-Betti numbers ([36]) $\beta_k^{(2)}(A_s(n))$ vanish for $k \geq 4$, and we have as well $\beta_0^{(2)}(A_s(n)) = 0$ by [37].

7 Hopf algebras with a projection

It is natural to ask whether similar results to those of Section 2 hold for Gerstenhaber-Schack cohomological dimension. A positive answer to Question 1.2 would of course provide an affirmative answer. So far, our only positive result in this direction is the following one, in the setting of Hopf algebras with a projection [47, 40].

Proposition 7.1. *Let* $B \subset A$ *be a Hopf subalgebra. Assume that there exists a Hopf algebra map* $\pi : A \to B$ *such that* $\pi_{|B} = \mathrm{id}_B$ *and that* A *is cosemisimple. Then we have* $\mathrm{cd}_{\mathrm{GS}}(B) \leq \mathrm{cd}_{\mathrm{GS}}(A)$*.*

Proof. The inclusion $B \subset A$ together with the Hopf algebra map $\pi : A \to B$ induce a vector space preserving linear exact tensor functor

$$F : \mathcal{YD}_A^A \longrightarrow \mathcal{YD}_B^B$$

where if V is Yetter-Drinfeld module over A, then $F(V) = V$ as a vector space, the B-module structure is the restriction of that of A, and the B-comodule structure is given by $(\mathrm{id}_V \otimes \pi)\alpha$, where α is the original co-action of A. We claim that it is enough to show that F sends (relative) projective Yetter-Drinfeld modules over A to (relative) projective Yetter-Drinfeld modules over B. Indeed, if we have a length n resolution of the trivial Yetter-Drinfeld module over A by (relative) projectives, the functor F will transform it into a a length n resolution of the trivial Yetter-Drinfeld module over B by (relative) projectives, and hence by Proposition 5.2, we have $\mathrm{cd}_{\mathrm{GS}}(B) \leq \mathrm{cd}_{\mathrm{GS}}(A)$.

As usual, put $R = {}^{\mathrm{co}B}A = \{a \in A \mid \pi(a_{(1)}) \otimes a_{(2)} = 1 \otimes a\}$. This is a subalgebra of A and we have $(\mathrm{id} \otimes \pi)\Delta(R) \subset R \otimes B$, which endows R with a right B-comodule structure. For any $a \in A$, we have $a_{(2)}\pi S^{-1}(a_{(1)}) \in R$ (since A is cosemisimple, its antipode is bijective), and thus we have a linear isomorphism [47, 40]

$$\begin{aligned} A &\longrightarrow R \otimes B \\ a &\longmapsto a_{(3)}\pi S^{-1}(a_{(2)}) \otimes \pi(a_{(1)}) \end{aligned}$$

whose inverse is the restriction of the multiplication of A. Let V be a right A-comodule: it also has a right B-comodule structure obtained using the projection $\pi : A \to B$, that we denote V_π. Consider now the map

$$\begin{aligned} F(V \boxtimes A) &\longrightarrow (V_\pi \otimes R) \boxtimes B \\ v \otimes a &\longmapsto v \otimes a_{(3)}\pi S^{-1}(a_{(2)}) \otimes \pi(a_{(1)}) \end{aligned}$$

This is an isomorphism by the previous considerations, and it is a direct verification to check that it is a morphism of Yetter-Drinfeld modules over B. Hence the functor F sends free Yetter-Drinfeld modules over A to free Yetter-Drinfeld modules over B, and since it is additive, it sends, by Proposition 4.2, projective Yetter-Drinfeld modules over A to projective Yetter-Drinfeld modules over B. This concludes the proof. $\square$

As an illustration, consider the hyperoctahedral Hopf algebra $A_h(n)$ [7]. This is the algebra presented by generators a_{ij}, $1 \leq i,j \leq n$, submitted to the relations

$$\sum_{l=1}^{n} a_{li}^2 = 1 = \sum_{l=1}^{n} a_{il}^2, \ a_{ik}a_{ij} = 0 = a_{ji}a_{ki} \text{ if } j \neq k, \ 1 \leq i,j,k \leq n$$

Its Hopf algebra structure is given by the same formulas as those for $A_s(n)$. There exist Hopf algebra maps $i : A_s(n) \to A_h(n)$, $x_{ij} \mapsto a_{ij}^2$, $\pi : A_h(n) \to A_s(n)$, $a_{ij} \mapsto x_{ij}$, such that $\pi i = \mathrm{id}$. Hence we deduce from the previous proposition that $\mathrm{cd}_{\mathrm{GS}}(A_h(n)) \geq \mathrm{cd}_{\mathrm{GS}}(A_s(n))$, and hence by Theorem 6.5, if $n \geq 4$, we have $\mathrm{cd}_{\mathrm{GS}}(A_h(n)) \geq \mathrm{cd}_{\mathrm{GS}}(A_s(n)) = 3$ (since $A_h(n)$ is cosemisimple of Kac type, this could be deduced as well from the combination of Proposition 3.1 and Corollary 5.10).

References

[1] N. Andruskiewitsch, J. Cuadra, On the structure of (co-Frobenius) Hopf algebras, *J. Noncommut. Geom.* 7 (2013), no. 1, 83-104.

[2] N. Andruskiewitsch, S. Dascalescu, Co-Frobenius Hopf algebras and the coradical filtration, *Math. Z.* 243 (2003), no. 1, 145-154.

[3] N. Andruskiewitsch, J. Devoto, Extensions of Hopf algebras, *St. Petersburg Math. J.* 7 (1996), no. 1, 17-52.

[4] S. Arkhipov, D. Gaitsgory, Another realization of the category of modules over the small quantum group, *Adv. Math.* 173 (2003), no. 1, 114-143.

[5] T. Banica, Symmetries of a generic coaction, *Math. Ann.* 314 (1999), no. 4, 763-780.

[6] T. Banica, Quantum groups and Fuss-Catalan algebras, *Comm. Math. Phys.* 226 (2002), no. 1, 221-232.

[7] T. Banica, J. Bichon, B. Collins, The hyperoctahedral quantum group, *J. Ramanujan Math. Soc.* 22 (2007), no. 4, 345-384.

[8] J. Bichon, Galois reconstruction of finite quantum groups, *J. Algebra* 230 (2000), no. 2, 683-693.

[9] J. Bichon, The representation category of the quantum group of a non-degenerate bilinear form, *Comm. Algebra* 31 (2003), no. 10, 4831-4851.

[10] J. Bichon, Hochschild homology of Hopf algebras and free Yetter-Drinfeld resolutions of the counit, *Compositio Math.* 149 (2013), no. 4, 658-678.

[11] J. Bichon, Hopf-Galois objects and cogroupoids, *Rev. Un. Mat. Argentina* 55 (2014), no. 2, 11-69.

[12] K.A. Brown, K.R Goodearl, Lectures on algebraic quantum groups, Advanced Courses in Mathematics CRM Barcelona, Birkhäuser Verlag, 2002.

[13] K.A. Brown, J.J. Zhang, Dualising complexes and twisted Hochschild (co)homology for Noetherian Hopf algebras, *J. Algebra* 320 (2008), no. 5, 1814-1850.

[14] K.S. Brown, Cohomology of groups, Graduate Texts in Mathematics 87, Springer-Verlag, New York-Berlin, 1982.

[15] S. Caenepeel, G. Militaru, S. Zhu, Crossed modules and Doi-Hopf modules, *Israel J. Math.* 100 (1997), 221-247.

[16] A. Chirvasitu, Cosemisimple Hopf algebras are faithfully flat over Hopf subalgebras, *Algebra Number Theory* 8 (2014), no. 5, 1179-1199.

[17] B. Collins, J. Härtel, A. Thom, Homology of free quantum groups. *C. R. Math. Acad. Sci. Paris* 347 (2009), no. 5-6, 271-276.

[18] A. Davydov, Twisting of monoidal structures, arXiv:q-alg/9703001.

[19] A. De Rijdt, N. Vander Vennet, Actions of monoidally equivalent compact quantum groups and applications to probabilistic boundaries, *Ann. Inst. Fourier* 60 (2010), no. 1, 169-216.

[20] M.S. Dijkhuizen, The double covering of the quantum group $SO_q(3)$, *Rend. Circ. Math. Palermo Serie II* 37 (1994), 47-57.

[21] M. Dubois-Violette, G. Launer, The quantum group of a non-degenerate bilinear form, *Phys. Lett. B* 245, no.2 (1990), 175-177.

[22] M.J. Dunwoody, Accessibility and groups of cohomological dimension one, *Proc. London Math. Soc.* 38 (1979), no. 2, 193-215.

[23] P. Etingof, S. Gelaki, On finite-dimensional semisimple and cosemisimple Hopf algebras in positive characteristic. *Internat. Math. Res. Notices* 1998, no. 16, 851-864.

[24] P. Etingof, S. Gelaki, D. Nikshych, V. Ostrik, Tensor categories. Mathematical Surveys and Monographs 205, American Mathematical Society, Providence, RI, 2015.

[25] P. Feng, B. Tsygan, Hochschild and cyclic homology of quantum groups. *Comm. Math. Phys.* 140 (1991), no. 3, 481-521.

[26] M. Gerhold, Additive deformations of Hopf algebras, *J. Algebra* 339 (2011), 114-134.

[27] M. Gerstenhaber, S. Schack, Bialgebra cohomology, deformations and quantum groups, *Proc. Nat. Acad. Sci. USA* 87 (1990), no. 1, 78-81.

[28] M. Gerstenhaber, S. Schack, Algebras, bialgebras, quantum groups, and algebraic deformations, *Contemp. Math.* 134 (1992), 51-92.

[29] V. Ginzburg, S. Kumar, Cohomology of quantum groups at roots of unity, *Duke Math. J.* 69 (1993), no. 1, 179-198.

[30] D.I. Gurevich, Algebraic aspects of the quantum Yang-Baxter equation, *Leningrad Math. J.* 2 (1991), no. 4, 801-828.

[31] T. Hadfield, U. Krähmer, Twisted homology of quantum SL(2), *K-Theory* 34 (2005), no. 4, 327-360.

[32] T. Hadfield, U. Krähmer, On the Hochschild homology of quantum SL(N), *C. R. Math. Acad. Sci. Paris* 343 (2006), no. 1, 9-13.

[33] L. Kadison, New examples of Frobenius extensions. University Lecture Series 14, American Mathematical Society, Providence, RI, 1999.

[34] C. Kassel, Quantum groups, GTM 155, Springer, 1995.

[35] A. Klimyk, K. Schmüdgen, Quantum groups and their representations, Texts and Monographs in Physics, Springer, 1997.

[36] D. Kyed, L^2-homology for compact quantum groups, *Math. Scand.* 103 (2008), no. 1, 111-129.

[37] D. Kyed, On the zeroth L^2-homology of a quantum group, *Münster J. Math.* 4 (2011), 119-128.

[38] R.G. Larson, D.E. Radford, Finite-dimensional cosemisimple Hopf algebras in characteristic 0 are semisimple, *J. Algebra* 117 (1988), no. 2, 267-289.

[39] B.I. Lin, Semiperfect coalgebras, *J. Algebra* 49 (1977), no. 2, 357-373.

[40] S. Majid, Cross products by braided groups and bosonization, *J. Algebra* 163 (1994), no. 1, 165-190.

[41] M. Mastnak, S. Witherspoon, Bialgebra cohomology, pointed Hopf algebras, and deformations, *J. Pure Appl. Algebra* 213 (2009), no. 7, 1399-1417.

[42] S. Montgomery, Hopf algebras and their actions on rings, Amer. Math. Soc. 1993.

[43] C. Mrozinski, Quantum automorphism groups and SO(3)-deformations, *J. Pure Appl. Algebra* 219 (2015), no. 1, 1-32.

[44] E. Müller, H.J. Schneider, Quantum homogeneous spaces with faithfully flat module structures, *Israel J. Math.* 111 (1999), 157-190.

[45] F. Panaite, D. Stefan, Deformation cohomology for Yetter-Drinfeld modules and Hopf (bi)modules. *Comm. Algebra* 30 (2002), no. 1, 331-345.

[46] B. Parshall, J. Wang, On bialgebra cohomology, *Bull. Soc. Math. Belg. Sér. A* 42 (1990), no. 3, 607-642.

[47] D.E. Radford, The structure of Hopf algebras with a projection, *J. Algebra* 92 (1985), no. 2, 322-347.

[48] D.E. Radford, Minimal quasitriangular Hopf algebras, *J. Algebra* 157 (1993), no. 2, 285-315.

[49] P. Schauenburg, Hopf modules and Yetter-Drinfeld modules, *J. Algebra* 169 (1994), no. 3, 874-890.

[50] P. Schauenburg, Hopf-Galois and bi-Galois extensions, *Fields Inst. Comm.* 43 (2004), 469-515.

[51] H.J. Schneider, Normal basis and transitivity of crossed products for Hopf algebras, *J. Algebra* 152 (1992), no. 2, 289-312.

[52] H.-J. Schneider, Some remarks on exact sequences of quantum groups, *Comm. Algebra* 21 (1993), no. 9, 3337-3357.

[53] S. Shnider, S. Sternberg, Quantum groups. From coalgebras to Drinfeld algebras. A guided tour. Graduate Texts in Mathematical Physics, II. International Press, Cambridge, MA, 1993.

[54] D. Stefan, Hochschild cohomology on Hopf Galois extensions, *J. Pure Appl. Algebra* 103 (1995), no. 2, 221-233.

[55] D. Stefan, The set of types of n-dimensional semisimple and cosemisimple Hopf algebras is finite, *J. Algebra* 193 (1997), no. 2, 571-580.

[56] R. Taillefer, Cohomology theories of Hopf bimodules and cup-product, *Algebr. Represent. Theory* 7 (2004), no. 5, 471-490.

[57] R. Taillefer, Injective Hopf bimodules, cohomologies of infinite dimensional Hopf algebras and graded-commutativity of the Yoneda product, *J. Algebra* 276 (2004), no. 1, 259-279.

[58] R. Taillefer, Bialgebra cohomology of the duals of a class of generalized Taft algebras, *Comm. Algebra* (2007), no. 4, 1415-1420.

[59] M. Takeuchi, Relative Hopf modules-equivalences and freeness criteria, *J. Algebra* 60 (1979), no. 2, 452-471.

[60] D.N. Yetter, Abelian categories of modules over a (lax) monoidal functor, *Adv. Math.* 174 (2003), no. 2, 266-309.

[61] S. Wang, Quantum symmetry groups of finite spaces, *Comm. Math. Phys.* 195 (1998), no. 1, 195-211.

[62] C. Weibel, An Introduction to Homological Algebra, Cambridge University Press, 1994.

Julien Bichon
Laboratoire de Mathématiques
Université Blaise Pascal
Campus des Cézeaux
3 place Vasarely
63178 Aubière Cedex
France
Julien.Bichon@math.univ-bpclermont.fr

A Characterization of Semiprojectivity for Subhomogeneous C^*-Algebras

Dominic Enders

Received: October 1, 2015
Revised: April 22, 2016

Communicated by Joachim Cuntz

Abstract. We study semiprojective, subhomogeneous C^*-algebras and give a detailed description of their structure. In particular, we find two characterizations of semiprojectivity for subhomogeneous C^*-algebras: one in terms of their primitive ideal spaces and one by means of special direct limit structures over one-dimensional NCCW complexes. These results are obtained by working out several new permanence results for semiprojectivity, including a complete description of its behavior with respect to extensions by homogeneous C^*-algebras.

2010 Mathematics Subject Classification: Primary 46L05; Secondary 46L80, 46L85, 54C55, 54F50
Keywords and Phrases: C^*-algebras, semiprojectivity, subhomogeneous, quantum permutation algebras

1 Introduction

The concept of semiprojectivity is a type of perturbation theory for C^*-algebras which has become a frequently used tool in many different aspects of C^*-algebra theory. Due to a certain kind of rigidity, semiprojective C^*-algebras are technically important in various situations. In particular, the existence and comparison of limit structures via approximate interwinings, which is an integral part of the Elliott classification program, often relies on perturbation properties of this type. This is one of the reasons why direct limits over

This work was supported by the SFB 878 *Groups, Geometry and Actions* and the Danish National Research Foundation through the *Centre for Symmetry and Deformation* (DNRF92).

semiprojective C^*-algebras, e.g., AF- or A$\mathbb{T}$-algebras, are particularly tractable and one therefore constructs models preferably from semiprojective building blocks. The most popular of those are without doubt the non-commutative CW-complexes (NCCWs) introduced by Eilers, Loring and Pedersen. These are in fact semiprojective in dimension one ([ELP98], but see also [End14]). In this paper, we study semiprojectivity for general subhomogeneous C^*-algebras and see whether there exist more interesting examples, i.e., besides the one-dimensional NCCW complexes (1-NCCWs), that could possibly serve as useful building blocks in the construction of ASH-algebras. In Theorem 5.1.2, we give two characterizations of semiprojectivity for subhomogenous C^*-algebras: an abstract one in terms of primitive ideal spaces and a concrete one by means of certain limit structures. These show that it is quite a restriction for a subhomogeneous C^*-algebra to be semiprojective, though many examples beyond the class of 1-NCCWs exist. On the other hand, a detailed study of the structure of these algebras further reveals that they can always be approximated by 1-NCCWs in a very strong sense, see Corollary 5.2.1, and hence essentially share the same properties.

The work of this paper is based on the characterization of semiprojectivity for commutative C^*-algebras, which was recently obtained by Sørensen and Thiel in [ST12]. They showed that a commutative C^*-algebra $\mathcal{C}(X)$ is semiprojective if and only if X is an absolute neighborhood retract of dimension at most 1 (a 1-ANR), thereby confirming a conjecture of Blackadar and generalizing earlier work of Chigogidze and Dranishnikov on the projective case ([CD10]). Their characterization further applies to trivially homogeneous C^*-algebras, i.e. to algebras of the form $\mathcal{C}(X, \mathbb{M}_n)$. In a first step, we generalize their result to general homogeneous C^*-algebras. The main difficulty, however, is to understand which ways of 'gluing together' several homogeneous C^*-algebras preserve semiprojectivity, or more precisely: Which extensions of semiprojective, homogeneous C^*-algebras are again semiprojective? Conversely, is semiprojectivity preserved when passing to a homogeneous subquotient? These questions essentially ask for the permanence behavior of semiprojectivity along extensions of the form $0 \to \mathcal{C}_0(X, \mathbb{M}_n) \to A \to B \to 0$. While it is known that the permanence properties of semiprojectivity with respect to extensions are rather bad in general, we are able to work out a complete description of its behavior in the special case of extensions by homogeneous ideals, see Theorem 4.3.2. With this permanence result at hand, it is then straightforward to characterize semiprojectivity for subhomogeneous C^*-algebras in terms of their primitive ideal spaces. In particular, it is a necessary condition that the subspaces corresponding to a fixed dimension are all 1-ANRs. Combining this with the structure result for one-dimensional ANR-spaces from [ST12], we further obtain a more concrete description of semiprojective, subhomogeneous C^*-algebras by identifying them with certain special direct limits of 1-NCCWs.

This paper is organized as follows. In section 2, we briefly recall some topological definitions and results that will be used troughout the paper. We further remind the reader of some facts about semiprojectivity, subhomogeneous C^*-algebras and their primitive ideal spaces. We then start by constructing a lifting problem which is unsolvable for strongly quasidiagonal C^*-algebras. This lifting problem then allows us to extend the results of [ST12] from the commutative to the homogeneous case.

Section 3 contains a number of new contructions for semiprojective C^*-algebras. We first introduce a technique to extend lifting problems, a method that can be used to show that in certain situations semiprojectivity passes to ideals. After that, we introduce a class of maps which give rise to direct limits that preserve semiprojectivity. Important examples of such maps are given and discussed.

Section 4 is devoted to the study of extensions by homogeneous C^*-algebras, i.e. extensions of the form $0 \to \mathcal{C}_0(X, \mathbb{M}_n) \to A \to B \to 0$. In 4.1, we define and study a certain set-valued retract map $R\colon \operatorname{Prim}(A) \to 2^{\operatorname{Prim}(B)}$ associated to such an extension. We discuss regularity concepts for R, i.e. continuity and finiteness conditions, and show how regularity of R relates to lifting properties of the corresponding Busby map and, by that, to splitting properties of the extension itself. In particular, we identify conditions under which regularity of R implies the existence of a splitting map $s\colon B \to A$ with good multiplicative properties. After that, we verify the required regularity properties for R in the case of a semiprojective extension A. In section 4.2 it is shown how certain limit structures for the space X give rise to limit structures for the extension A, again provided that the associated retract map R is sufficiently regular. Putting all these results together in 4.3, we find a '2 out of 3'-type statement, Theorem 4.3.2, which gives a complete description for the behavior of semiprojectivity along extensions of the considered type.

In section 5.1, we use this permanence result to work out two characterizations of semiprojectivity for subhomogeneous C^*-algebras. These are presented in Theorem 5.1.2, the main result of this paper. Based on this, we find a number of consequences for the structure of these algebras, e.g. information about their K-theory and dimension. Further applications, such as closure and approximation properties, are discussed in 5.2. We finish by illustrating how this also gives a simple method to exclude semiprojectivity and show that the higher quantum permutation algebras are not semiprojective.

2 Preliminaries

2.1 The structure of 1-dimensional ANR-spaces

We are particularly interested in ANR-spaces of dimension at most one. The structure of these spaces has been studied and described in detail in [ST12, section 4]. Here we recall the most important notions and results. More information about ANR-spaces can be found in [Bor67]. For proofs and further reading on the theory of continua, we refer the reader to Nadler's book [Nad92].

Definition 2.1.1. *A compact, metric space X is an absolute retract (abbreviated AR-space) if every map $f\colon Z \to X$ from a closed subspace Z of a compact,* *metric space Y extends to a map $g\colon Y \to X$, i.e. $g \circ \iota = f$ with $\iota\colon Z \to Y$ the inclusion map:*

$$\begin{tikzcd} & Y \arrow[dl, dashed, "g"'] \\ X & Z \arrow[l, "f"] \arrow[u, "\iota"'] \end{tikzcd}$$

If every map $f\colon Z \to X$ from a closed subspace Z of a compact, metric space Y extends to a map $g\colon V \to X$ on a closed neighborhood V of Z

$$\begin{tikzcd} & Y \\ & V \arrow[u] \arrow[dl, dashed, "g"'] \\ X & Z \arrow[l, "f"] \arrow[u, "\iota"'] \end{tikzcd}$$

then X is called an absolute neighborhood retract (abbreviated ANR-space).

A compact, locally connected, metric space is called a Peano space. A connected Peano space is called a Peano continuum. By the Hahn-Mazurkiewicz Theorem, these continua can be characterized as the continuous images of the unit interval. In particular, every Peano continuum is path-connected, while every Peano space is locally path-connected.
Now given an ANR-space X, we can embed it into the Hilbert cube $\mathcal{Q}$ and obtain a retract from a neighborhood of X in $\mathcal{Q}$ onto X. Hence an ANR-space inherits all local properties of the Hilbert cube which are preserved under retracts. These properties include local connectedness, so that all ANR-spaces are Peano spaces. The converse, however, is not true in general. But as we will see, it is possible to identify the ANR-spaces among all Peano spaces, at least in the one-dimensional case.
A closed subspace Y of a space X is a retract of X if there exists a continuous map $r\colon X \to Y$ such that $r_{|Y} = \mathrm{id}_Y$. If the retract map $r\colon X \to Y$ regarded as a map to X is homotopic to the identity, then Y is called a deformation retract of X. It is a strong deformation retract if in addition the homotopy can be chosen to fix the subspace Y. The following concept of a core continuum is due to Meilstrup. It is crucial for understanding the structure of one-dimensional ANR-spaces.

DEFINITION + LEMMA 2.1.2 ([Mei05]). *Let X be a non-contractible one-dimensional Peano continuum. Then there exists a unique strong deformation retract which contains no further proper deformation retract. We call it the core of X and denote it by* $\mathrm{core}(X)$.

As in [ST12], we define the core of a contractible, one-dimensional Peano continuum to be any fixed point. Many questions about one-dimensional Peano

continua can be reduced to questions about their cores. This reduction step uses a special retract map, the so-called first point map.
Recall that an arc between two points $x_0, x_1 \in X$ is a path $[0,1] \to X$ from x_0 to x_1 which is a homeomorphism onto its image.

DEFINITION + LEMMA 2.1.3 ([ST12, 4.14-16]). *Let X be a one-dimensional Peano continuum and Y a subcontinuum with $\mathrm{core}(X) \subset Y$. For each $x \in X\backslash Y$ there is a unique point $r(x) \in Y$ such that $r(x)$ is a point of an arc in X from x to any point of Y. Setting $r(x) = x$ for all $x \in Y$, we obtain a map $r\colon X \to Y$. This map is called the first point map, it is continuous and a strong deformation retract from X onto Y.*

The following follows directly from the proof of [ST12, Lemma 4.14].

LEMMA 2.1.4. *Let X be a one-dimensional Peano continuum, $Y \subseteq X$ a subcontinuum containing $\mathrm{core}(X)$ and $r\colon X \to Y$ the first point map onto Y. Then the following is true:*

(i) For every point $x \in X\backslash Y$ there exists an arc from x to $r(x) \in Y$ which is unique up to reparametrization.

(ii) If α is a path from $x \in X\backslash Y$ to $y \in Y$, then $r(\mathrm{im}(\alpha)) \subseteq \mathrm{im}(\alpha)$.

The simplest example of a one-dimensional Peano space is a graph, i.e. a finite, one-dimensional CW-complex. The order of a point x in a graph X is defined as the smallest number $n \in \mathbb{N}$ such that for every neighborhood V of x there exists an open neighborhood $U \subseteq V$ of x with $|\partial U| = |\overline{U}\backslash U| \leq n$. We denote the order of x in X by $\mathrm{order}(x, X)$.
Given a one-dimensional Peano continuum X, one can reconstruct the space X from its core by 'adding' the arcs which connect points of $X\setminus \mathrm{core}(X)$ with the core as described in 2.1.4. This procedure yields a limit structure for one-dimensional Peano spaces which first appeared as Theorem 4.17 of [ST12]. In the case of one-dimensional ANR-spaces, the core is a finite graph and hence the limit structure entirely consists of finite graphs.

THEOREM 2.1.5 ([ST12, Theorem 4.17]). *Let X be a one-dimensional Peano continuum. Then there exists a sequence $\{Y_k\}_{k=1}^{\infty}$ such that*

(i) each Y_k is a subcontinuum of X.

(ii) $Y_k \subset Y_{k+1}$.

(iii) $\lim_k Y_k = X$.

(iv) $Y_1 = \mathrm{core}(X)$ and for each k, Y_{k+1} is obtained from Y_k by attaching a line segment at a single point, i.e., $\overline{Y_{k+1}\backslash Y_k}$ is an arc with end point p_k such that $\overline{Y_{k+1}\backslash Y_k} \cap Y_k = \{p_k\}$.

(v) letting $r_k\colon X \to Y_k$ be the first point map for Y_k we have that $\{r_k\}_{k=1}^{\infty}$ converges uniformly to the identity map on X.

If X is also an ANR, then all Y_k are finite graphs. If X is even contractible (i.e. an AR), then $\mathrm{core}(X)$ *is just some point and all Y_k are finite trees.*

We will need a local criterion for identifying one-dimensional ANR-spaces among general Peano spaces. It was observed by Ward how to get such a characterization in terms of embeddings of circles.

Definition 2.1.6. *Let X be a compact, metric space, then X does not contain small circles if there is an $\epsilon > 0$ such that* $\mathrm{diam}(\iota(S^1)) \geq \epsilon$ *for every embedding* $\iota \colon S^1 \to X$.

Note that the property of containing arbitrarily small circles does not depend on the particular choice of metric.

Theorem 2.1.7 ([War60]). *For a Peano space X the following are equivalent:*

(i) X does not contain small circles.

(ii) X is an ANR-space of dimension at most one.

This statement can also be interpreted as follows. Non-embeddability of circles into X is the same as uniqueness of arcs in X, i.e. an arc between to two given points is unique up to reparametrization. More precisely, a Peano continuum is a one-dimensional AR-space if and only if there is no embedding $S^1 \hookrightarrow X$ if and only if X has unique arcs. Similarly, Theorem 2.1.7 can be read as: A Peano continuum X is a one-dimensional ANR-space if and only if it has locally unique arcs, meaning that every point has a neighbouhood in which any two points can be joined by a unique arc.

2.2 Subhomogeneous C^*-algebras

In this section we collect some well known results on subhomogeneous C^*-algebras. In particular, we recall some facts on their primitive ideal spaces. More detailed information can be found in [Dix77, Chapter 3] and [Bla06, Section IV.1.4].

Definition 2.2.1. *Let $N \in \mathbb{N}$. A C^*-algebra A is N-homogeneous if all its irreducible representations are of dimension N. A is N-subhomogeneous if every irreducible representation of A has dimension at most N.*

The standard example of a N-homogeneous C^*-algebra is $\mathcal{C}_0(X, \mathbb{M}_N)$ for some locally compact space X. As the next proposition shows, subhomogeneous C^*-algebras can be characterized as subalgebras of such. A proof of this fact can be found in [Bla06, IV.1.4.3-4].

Proposition 2.2.2. *A C^*-algebra A is N-subhomogeneous if and only if it is isomorphic to a subalgebra of some N-homogeneous C^*-algebra $\mathcal{C}(X, \mathbb{M}_N)$. If A is separable, we may choose X to be the Cantor set K.*

Example 2.2.3 (1-NCCWs). *One of the most important examples of subhomogeneous C^*-algebras is the class of non-commutative CW-complexes (NCCWs) defined by Eilers, Loring and Pedersen in [ELP98]. The one-dimensional NCCWs, which we will abbreviate by 1-NCCWs, are defined as pullbacks of the form*

$$\begin{tikzcd} \text{1-NCCW} \arrow[r, dashed] \arrow[d, dashed] & G \arrow[d] \\ \mathcal{C}([0,1],F) \arrow[r, "\mathrm{ev}_0 \oplus \mathrm{ev}_1"] & F \oplus F \end{tikzcd}$$

with F and G finite-dimensional C^*-algebras. These are particularly interesting since they are semiprojective by [ELP98, Theorem 6.2.2].

For a subhomogeneous C^*-algebra A, the primitive ideal space $\mathrm{Prim}(A)$, i.e. the set of kernels of irreducible representations endowed with the Jacobson topology, contains a lot of information. Another useful decription of the topology on $\mathrm{Prim}(A)$ is given by the folllowing lemma which we will make use of regularly. For an ideal J in a C^*-algebra A we write $\|x\|_J$ to denote the norm of the image of the element $x \in A$ in the quotient A/J.

Lemma 2.2.4 ([Bla06, II.6.5.6]). *Let A be a C^*-algebra.*

1. *If $x \in A$, define $\check{x}\colon \mathrm{Prim}(A) \to \mathbb{R}_{\geq 0}$ by $\check{x}(J) = \|x\|_J$. Then $\check{x}$ is lower semicontinuous.*
2. *If $\{x_i\}$ is a dense set in the unit ball of A, and $U_i = \{J \in \mathrm{Prim}(A) : \check{x}_i(J) > 1/2\}$, then $\{U_i\}$ forms a base for the topology of $\mathrm{Prim}(A)$.*
3. *If $x \in A$ and $\lambda > 0$, then $\{J \in \mathrm{Prim}(A)\colon \check{x}(J) \geq \lambda\}$ is compact (but not necessarily closed) in $\mathrm{Prim}(A)$.*

Since we will mostly be interested in finite-dimensional representations, we consider the subspaces

$$\mathrm{Prim}_n(A) = \{\ker(\pi) \in \mathrm{Prim}(A)\colon \dim(\pi) = n\}$$

for each finite n. Similarly, we write

$$\mathrm{Prim}_{\leq n}(A) = \{\ker(\pi) \in \mathrm{Prim}(A)\colon \dim(\pi) \leq n\} = \bigcup_{k \leq n} \mathrm{Prim}_k(A).$$

The following theorem describes the structure of these subspaces of $\mathrm{Prim}(A)$ and the relations between them.

Theorem 2.2.5 ([Dix77, 3.6.3-4]). *Let A be a C^*-algebra. The following holds for each $n \in \mathbb{N}$:*

(i) $\mathrm{Prim}_{\leq n}(A)$ *is closed in* $\mathrm{Prim}(A)$.

(ii) $\mathrm{Prim}_n(A)$ *is open in* $\mathrm{Prim}_{\leq n}(A)$.

(iii) $\mathrm{Prim}_n(A)$ *is locally compact and Hausdorff.*

Now assume that A is a N-subhomogeneous C^*-algebra. In this case Theorem 2.2.5 gives a set-theoretical (but in general not a topological) decomposition of its primitive spectrum

$$\mathrm{Prim}(A) = \bigsqcup_{n=1}^{N} \mathrm{Prim}_n(A).$$

While each subspace in this decomposition is nice, in the sense that it is Hausdorff, $\mathrm{Prim}(A)$ itself is typically non-Hausdorff. In the subhomogeneous setting it is at least a T_1-space, i.e. points are closed. If we further assume A to be separable and unital, the space $\mathrm{Prim}(A)$ will also be separable and quasi-compact. Given a general C^*-algebra A, there is a one-to-one correspondence between (closed) ideals J of A and closed subsets of $\mathrm{Prim}(A)$. More precisely, one can identify $\mathrm{Prim}(A/J)$ with the closed subset $\{K \in \mathrm{Prim}(A)\colon J \subseteq K\}$. In particular, we can consider the quotient $A_{\leq n}$ corresponding to the closed subset $\mathrm{Prim}_{\leq n}(A) \subseteq \mathrm{Prim}(A)$. This quotient is the maximal n-subhomogeneous quotient of A and has the following universal property: Any *-homomorphism $\varphi\colon A \to B$ to some n-subhomogeneous C^*-algebra B factors uniquely through $A_{\leq n}$:

$$\begin{tikzcd} A \arrow[rr, "\varphi"] \arrow[dr] & & B \\ & A_{\leq n} \arrow[ur, dashed] & \end{tikzcd}$$

2.3 Semiprojective C^*-algebras

We recall the definition of semiprojectivity for C^*-algebras, the main property of study in this paper. More detailed information about lifting properties for C^*-algebras can be found in Loring's book [Lor97].

Definition 2.3.1 ([Bla85, Definition 2.10]). *A separable C^*-algebra A is semiprojective if for every C^*-algebra B and every increasing chain of ideals J_n in B with $J_\infty = \overline{\bigcup_n J_n}$, and for every *-homomorphism $\varphi\colon A \to B/J_\infty$ there exist $n \in \mathbb{N}$ and a *-homomorphism $\overline{\varphi}\colon A \to B/J_n$ making the following diagram commute:*

$$\begin{tikzcd} & B \arrow[d, "\pi_0^n"] \\ & B/J_n \arrow[d, "\pi_n^\infty"] \\ A \arrow[ur, dashed, "\overline{\varphi}"] \arrow[r, "\varphi"] & B/J_\infty \end{tikzcd}$$

In this situation, the map $\overline{\varphi}$ is called a partial lift of φ. The C^*-algebra A is projective if, in the situation above, we can always find a lift $\overline{\varphi}\colon A \to B$ for φ. Let $\mathcal{C}$ be a class of C^*-algebras. A C^*-algebra A is (semi)projective with respect to $\mathcal{C}$ if it satisfies the definitions above with the restriction that the C^*-algebras $B, B/J_n$ and B/J_∞ all belong to the class $\mathcal{C}$.

Remark 2.3.2. *One may also define semiprojectivity as a lifting property for maps to certain direct limits: an increasing sequence of ideals J_n in B gives an inductive system $(B/J_n)_n$ with surjective connecting maps $\pi_n^{n+1}\colon B/J_n \to B/J_{n+1}$ and limit (isomorphic to) B/J_∞. On the other hand, it is easily seen that every such system gives an increasing chain of ideals $(\ker(\pi_0^n))_n$. Hence, semiprojectivity is equivalent to being able to lift maps to $\varinjlim D_n$ to a finite stage D_n provided that all connecting maps of the system are surjective. It is sometimes more convenient to work in this picture.*

2.3.1 An unsolvable lifting problem

In order to show that a C^*-algebra does not have a certain lifting property, we need to construct unsolvable lifting problems. One such construction by Loring ([Lor97, Proposition 10.1.8]) uses the fact that normal elements in quotient C^*-algebras do not admit normal preimages in general, e.g. Fredholm operators of non-zero index. Here, we generalize Loring's construction and obtain a version which also works for almost normal elements. Combining this with Lin's theorem on almost normal matrices, we are able to construct unsolvable lifting problems not only for commutative C^*-algebras, as in Loring's case, but for the much larger class of strongly quasidiagonal C^*-algebras.

First we observe that almost normal elements in quotient C^*-algebras always admit (almost as) almost normal preimages. Given an element x of some C^*-algebra and $\epsilon > 0$, we say that x is ϵ-normal if $\|x^*x - xx^*\| \leq \epsilon\|x\|$ holds.

Lemma 2.3.3. *Let A, B be C^*-algebras and $\pi\colon A \to B$ a surjective *-homomorphism. Then for every ϵ-normal element $y \in B$ there exists a (2ϵ)-normal element $x \in A$ with $\pi(x) = y$ and $\|x\| = \|y\|$.*

Proof. Let $(u_\lambda)_{\lambda\in\Lambda}$ denote an approximate unit for $\ker(\pi)$ which is quasicentral for A. Pick any preimage x_0 of x with $\|x_0\| = \|x\|$ and set $x := (1 - u_{\lambda_0})x_0$ for a suitable $\lambda_0 \in \Lambda$. □

The next lemma is due to Halmos. A short proof using the Fredholm alternative can be found in [BH74, Lemma 2].

Lemma 2.3.4 (Halmos). *Let $S \in \mathcal{B}(H)$ be a proper isometry, then*

$$\operatorname{dist}(S, \{N + K \mid N, K \in \mathcal{B}(H),\ N \text{ normal},\ K \text{ compact}\}) = 1.$$

It is a famous result by H. Lin that in matrix algebras almost normal elements are uniformly close to normal ones ([Lin97]). A short, alternative proof involving semiprojectivity arguments can be found in [FR01].

Theorem 2.3.5 (Lin). *For every $\epsilon > 0$, there is a $\delta > 0$ so that, for any d and any X in $\mathbb{M}_d$ satisfying*

$$\|XX^* - X^*X\| \leq \delta \quad \textit{and} \quad \|X\| \leq 1$$

there is a normal Y in $\mathbb{M}_d$ such that

$$\|X - Y\| \leq \epsilon.$$

The following is the basis for most of our unsolvable lifting problems appearing in this paper. Recall that a C^*-algebra A is strongly quasidiagonal if every representation of A is quasidiagonal. See [Bla06, Section V.4.2] or [Bro00] for more information on quasidiagonality.
In the following, let $\mathcal{T}$ denote the Toeplitz algebra $C^*(S|S^*S = 1)$ and $\varrho\colon \mathcal{T} \to \mathcal{C}(S^1)$ the quotient map given by mapping S to the canonical generator z of $\mathcal{C}(S^1)$.

Proposition 2.3.6. *There exists $\delta > 0$ such that the following holds for all $n \in \mathbb{N}$: If A is strongly quasidiagonal and $\varphi\colon A \to \mathcal{C}(S^1) \otimes \mathbb{M}_n$ is any *-homomorphism with $\operatorname{dist}(z \otimes 1_n, \operatorname{im}(\varphi)) < \delta$, then φ does not lift to a *-homomorphism from A to $\mathcal{T} \otimes \mathbb{M}_n$:*

$$\begin{tikzcd} & \mathcal{T} \otimes \mathbb{M}_n \arrow[d, "\varrho\otimes\mathrm{id}"] \\ A \arrow[ur, dashed, "\nexists"] \arrow[r, "\varphi"] & \mathcal{C}(S^1) \otimes \mathbb{M}_n \end{tikzcd}$$

Proof. Choose $\delta' > 0$ corresponding to $\epsilon = 1/6$ as in Theorem 2.3.5 and set $\delta = \delta'/14$. Let $a' \in A$ be such that $\|\varphi(a') - z \otimes 1_n\| < \delta$, then $\|[\varphi(a'), \varphi(a')^*]\| \leq 2\delta(\|\varphi(a')\| + 1) < 5\delta\|\varphi(a')\|$. Hence by Lemma 2.3.3 there exists a (10δ)-normal element $a \in A$ with $\varphi(a) = \varphi(a')$ and $5/6 < \|a\| = \|\varphi(a')\| < 6/5$. Now if ψ is a *-homomorphism with $(\varrho \otimes \mathrm{id}) \circ \psi = \varphi$ as indicated, we regard ψ as a representation on $\mathcal{H}^{\oplus n}$ with $\mathcal{T}$ generated by the unilateral shift S on $\mathcal{H}$. By assumption, ψ is then a quasidiagonal representation. In particular, $\psi(a)$ can be approximated arbitrarily well by block-diagonal operators ([Bro00, Theorem 5.2]). We may therefore choose a (11δ)-normal block-diagonal operator B with $5/6 \leq \|B\| \leq 6/5$ within distance at most $1/3$ from $\psi(a)$. Applying Lin's Theorem to the normalized, (14δ)-normal block-diagonal operator $\|B\|^{-1}B$ shows the existence of a normal element $N \in \mathcal{H}^{\oplus n}$ with $\|\psi(a) - N\| \leq 2/3$. But then we find

$$\begin{aligned} & \|(N - S \otimes 1_n) + \mathcal{K}(\mathcal{H}^{\oplus n})\| \\ \leq \; & \|N - \psi(a)\| + \|(\varrho \otimes \mathrm{id})(\psi(a) - S \otimes 1_n)\| \\ \leq \; & \tfrac{2}{3} + \|\varphi(a') - z \otimes 1_n\| \\ \leq \; & \tfrac{2}{3} + \delta < 1 \end{aligned}$$

in contradiction to Lemma 2.3.4. ∎

2.3.2 The homogeneous case

In [ST12], A. Sørensen and H. Thiel characterized semiprojectivity for commutative C^*-algebras. Moreover, they gave a description of semiprojectivity for homogeneous trivial fields, i.e. C^*-algebras of the form $\mathcal{C}_0(X, \mathbb{M}_N)$. Note that the projective case was settled earlier by A. Chigogidze and A. Dranishnikov in [CD10]. Their result is as follows.

THEOREM 2.3.7 ([ST12]). *Let X be a locally compact, metric space and $N \in \mathbb{N}$. Then the following are equivalent:*

1. *$\mathcal{C}_0(X, \mathbb{M}_N)$ is (semi)projective.*
2. *The one-point compactification αX is an A(N)R-space and* $\dim(X) \leq 1$.

The work of Sørensen and Thiel will be the starting point for our analysis of semiprojectivity for subhomogeneous C^*-algebras. In this section, we reduce the general N-homogeneous case to their result by showing that semiprojectivity for homogeneous, locally trivial fields implies global triviality. We further obtain some information about parts of the primitive ideal space for general semiprojective C^*-algebras.

LEMMA 2.3.8. *Let I be a N-homogeneous ideal in a C^*-algebra A. If A is semiprojective with respect to N-subhomogeneous C^*-algebras, then the one-point compactification $\alpha \operatorname{Prim}(I)$ is a Peano space. If A is semiprojective, we further have* $\dim(\alpha \operatorname{Prim}(I)) \leq 1$.

Proof. Let $A_{\leq N}$ be the maximal N-subhomogeneous quotient of A, then I is also an ideal in $A_{\leq N}$. Being N-homogeneous, the ideal I is isomorphic to the section algebra $\Gamma_0(E)$ of a locally trivial $\mathbb{M}_N$-bundle E over the locally compact, second countable, metrizable Hausdorff space $\operatorname{Prim}(I)$ by [Fel61, Theorem 3.2]. Since $A_{\leq N}$ is separable and N-subhomogenous, we can embed it into $\mathcal{C}(K, \mathbb{M}_N)$ with K the Cantor set by Proposition 2.2.2. Using the well known middle-third construction of $K = \varprojlim_k (\bigsqcup^{2^k} [0,1])$, we can write $\mathcal{C}(K)$ as a direct limit $\varinjlim_k \mathcal{C}([0,1])^{\oplus 2^k}$ with surjective connecting maps. After tensoring with $\mathbb{M}_N$, we obtain a lifting problem for A and hence can apply semiprojectivity of A with respect to N-subhomogenous C^*-algebras. As the solution to this lifting problem factors through $A_{\leq N}$, we obtain an embedding of $A_{\leq N}$ into $\mathcal{C}([0,1], \mathbb{M}_N)^{\oplus 2^k}$ for some k.

$$\begin{tikzcd}
 & A \arrow[d] \arrow[r, dashed] & \mathcal{C}([0,1], \mathbb{M}_N)^{\oplus 2^k} \arrow[d] \\
I \arrow[ur] \arrow[r] & A_{\leq N} \arrow[r] \arrow[ur, dotted] & \mathcal{C}(K, \mathbb{M}_N)
\end{tikzcd}$$

The restriction of this embedding to I induces a continuous surjection π of $\bigsqcup^{2^k} [0,1]$ onto $\alpha \operatorname{Prim}(I)$. By the Hahn-Mazurkiewicz Theorem ([Nad92, Theorem 8.18]), this shows that $\alpha \operatorname{Prim}(I)$ is a Peano space. Furthermore, we find a basis of compact neighborhoods consisting of Peano continua for any point x of $\alpha\,\mathrm{Prim}(I)$ by [Nad92, Theorem 8.10].

Now let A be semiprojective and assume that $\dim(\mathrm{Prim}(I)) = \dim(\alpha\,\mathrm{Prim}(I)) > 1$. Arguing precisely as in [ST12, Proposition 3.1], we use our basis of neighborhoods for points of $\mathrm{Prim}(I)$ to find arbitrarily small circles around a point $x \in \mathrm{Prim}(I)$. Using triviality of E around x, we obtain a lifting problem for A:

$$\begin{tikzcd} & A \arrow[r, dashed] \arrow[d] & \Big(\big(\bigoplus_{\mathbb{N}} \mathcal{T}\big)^+ / \big(\bigoplus_1^n \mathbb{K}\big)\Big) \otimes \mathbb{M}_N \arrow[d] \\ I \arrow[ur, "\subseteq"] \arrow[r] & \big(\bigoplus_{\mathbb{N}} \mathcal{C}(S^1)\big)^+ \otimes \mathbb{M}_N \arrow[r, equal] & \Big(\big(\bigoplus_{\mathbb{N}} \mathcal{T}\big)^+ / \big(\bigoplus_{\mathbb{N}} \mathbb{K}\big)\Big) \otimes \mathbb{M}_N \end{tikzcd}$$

Semiprojectivity of A allows us to solve this lifting problem. Now restrict a partial lift to the ideal I and consider its coordinates to obtain a commutative diagram

$$\begin{tikzcd} & \mathcal{T} \otimes \mathbb{M}_N \arrow[d] \\ I \arrow[ur, dashed] \arrow[r, two heads] & \mathcal{C}(S^1) \otimes \mathbb{M}_N. \end{tikzcd}$$

The map on the bottom is surjective since it is induced by the inclusion of one of the circles around x. But a diagram like this does not exist by Proposition 2.3.6 because I is homogeneous and by that strongly quasidiagonal. $\square$

COROLLARY 2.3.9. *Let A be a semiprojective C^*-algebra, then $\alpha\,\mathrm{Prim}_n(A)$ is a Peano space for every $n \in \mathbb{N}$.*

Proof. If A is semiprojective, each $A_{\leq n}$ is semiprojective with respect to n-subhomogeneous C^*-algebras. Hence we can apply Lemma 2.3.8 to the n-homogeneous ideal $\ker(A_{\leq n} \to A_{\leq n-1})$ in $A_{\leq n}$ whose primitive ideal space is homeomorphic to $\mathrm{Prim}_n(A)$. $\square$

It is known to the experts that there are no non-trivial $\mathbb{M}_n$-valued fields over one-dimensional spaces and we are indebted to L. Robert for pointing this fact out to us. Since we couldn't find a proof in the literature, we include one here.

LEMMA 2.3.10. *Let E be a locally trivial field of C^*-algebras over a separable, metrizable, locally compact Hausdorff space X with fiber $\mathbb{M}_N$ and $\Gamma_0(E)$ the corresponding section algebra. If $\dim(X) \leq 1$, then $\Gamma_0(E)$ is $\mathcal{C}_0(X)$-isomorphic to $\mathcal{C}_0(X, \mathbb{M}_N)$.*

Proof. First assume that X is compact. One-dimensionality of X implies that that the Dixmier-Douady invariant $\delta \in \check{H}^3(X, \mathbb{Z})$ corresponding to $\Gamma_0(E)$ vanishes. Therefore $\Gamma_0(E)$ is stably $\mathcal{C}(X)$-isomorphic to $\mathcal{C}(X, \mathbb{M}_N)$ by Dixmier-Douady classification (see e.g. [RW98, Corollary 5.56]). Let $\psi\colon \Gamma(E) \otimes \mathcal{K} \to$

$\mathcal{C}(X,\mathbb{M}_N)\otimes\mathcal{K}$ be such an isomorphism and note that $\Gamma(E)\cong \mathrm{her}(\psi(1_{\Gamma(E)}\otimes e))$ via ψ with e a minimal projection in $\mathcal{K}$. Equivalence of projections over one-dimensional spaces is completely determined by their rank by [Phi07, Proposition 4.2]. Since $\psi(1_{\Gamma(E)}\otimes e)$ and $1_{\mathcal{C}(X,\mathbb{M}_N)}\otimes e$ share the same rank N everywhere we therefore find $v\in\mathcal{C}(X,\mathbb{M}_N)\otimes\mathcal{K}$ with $v^*v=\psi(1_{\Gamma(E)}\otimes e)$ and $vv^*=1_{\mathcal{C}(X,\mathbb{M}_N)}\otimes e$. But then $\mathrm{Ad}(v)$ gives a $C(X)$-isomorphism from $\mathrm{her}(\psi(1_{\Gamma(E)}\otimes e))$ onto $\mathrm{her}(1_{\mathcal{C}(X,\mathbb{M}_N)}\otimes e)=\mathcal{C}(X,\mathbb{M}_N)$.
Now consider the case of non-compact X. Since X is σ-compact, it clearly suffices to prove the following: Given compact subsets $X_1\subseteq X_2$ of X and a $\mathcal{C}(X_1)$-isomorphism $\varphi_1\colon\Gamma(E_{|X_1})\to\mathcal{C}(X_1,\mathbb{M}_N)$ there exists a $\mathcal{C}(X_2)$-isomorphism $\varphi_2\colon\Gamma(E_{|X_2})\to\mathcal{C}(X_2,\mathbb{M}_N)$ extending φ_1. By the first part of the proof there is a $\mathcal{C}(X_2)$-isomorphism $\psi_2\colon\Gamma(E_{|X_2})\to\mathcal{C}(X_2,\mathbb{M}_N)$. One-dimensionality of X_1 implies $\check{H}^2(X_1,\mathbb{Z})=0$, which means that every $\mathcal{C}(X_1)$-automorphism of $\mathcal{C}(X_1,\mathbb{M}_N)$ is inner by [RW98, Theorem 5.42]. In particular, $\varphi_1\circ(\psi_2^{-1})_{|X_1}$ is of the form $\mathrm{Ad}(u)$ for some unitary $u\in\mathcal{C}(X_1,\mathbb{M}_N)$. It remains to extend u to a unitary in $\mathcal{C}(X_2,\mathbb{M}_N)$. This, however, follows from one-dimensionality of X and [HW48, Theorem VI.4]. ☐

We are now able to extend the results of [ST12] to the case of general N-homogeneous C^*-algebras:

Theorem 2.3.11. *Let A be a N-homogeneous C^*-algebra. The following are equivalent:*

1. *A is (semi)projective.*
2. *$A\cong\mathcal{C}_0(\mathrm{Prim}(A),\mathbb{M}_N)$ and $\alpha\,\mathrm{Prim}(A)$ is an A(N)R-space of dimension at most 1.*

Proof. We know that (1) implies $A\cong\mathcal{C}_0(\mathrm{Prim}(A),\mathbb{M}_N)$ by Lemma 2.3.8 and Lemma 2.3.10. The remaining implications are given by Theorem 2.3.7. ☐

3 Constructions for semiprojective C^*-algebras

Unfortunately, the class of semiprojective C^*-algebras lacks good permanence properties. In fact, semiprojectivity is not preserved by most C^*-algebraic standard constructions and the list of positive permanence results, most of which can be found in [Lor97], is surprisingly short. Here, we extend this list by a few new results.

3.1 Extending lifting problems

In this section, we introduce a technique to extend lifting problems from ideals to larger C^*-algebras. This technique can be used to show that in many situations lifting properties of a C^*-algebra pass to its ideals.

LEMMA 3.1.1. *Given a surjective inductive system of short exact sequences*

$$\begin{tikzcd}
0 \arrow[r] & C_n \arrow[r, "\iota_n"] \arrow[d, two heads, "\pi_n^{n+1}"'] & D_n \arrow[r, "\varrho_n"] \arrow[d, two heads, "\overline{\pi}_n^{n+1}"'] & E_n \arrow[r] \arrow[d, two heads, "\overline{\overline{\pi}}_n^{n+1}"'] & 0 \\
0 \arrow[r] & C_{n+1} \arrow[r, "\iota_{n+1}"] & D_{n+1} \arrow[r, "\varrho_{n+1}"] & E_{n+1} \arrow[r] & 0
\end{tikzcd}$$

and a commutative diagram of extensions

$$\begin{tikzcd}
0 \arrow[r] & \varinjlim C_n \arrow[r, "\iota_\infty"] & \varinjlim D_n \arrow[r, "\varrho_\infty"] & \varinjlim E_n \arrow[r] & 0 \\
0 \arrow[r] & I \arrow[r, "i"] \arrow[u, "\varphi"] & A \arrow[r, "p"] \arrow[u, "\overline{\varphi}"] & B \arrow[r] \arrow[u, "\overline{\overline{\varphi}}"] & 0
\end{tikzcd}$$

the following holds: If both A and B are semiprojective, then φ lifts to C_n for some n. If both A and B are projective, then φ lifts to C_1.

Proof. First observe that we may assume the *-homomorphism $\overline{\overline{\varphi}}$ to be injective since otherwise we simply pass to the system of extensions

$$\begin{tikzcd}
0 \arrow[r] & C_n \arrow[r, "\iota_n"] & D_n \oplus B \arrow[r, "\varrho_n \oplus \mathrm{id}_B"] & E_n \oplus B \arrow[r] & 0
\end{tikzcd}$$

and replace $\overline{\varphi}$ by $\overline{\varphi} \oplus p$ and $\overline{\overline{\varphi}}$ by $\overline{\overline{\varphi}} \oplus \mathrm{id}_B$. Using semiprojectivity of B, we can find a partial lift $\psi\colon B \to E_{n_0}$ of $\overline{\overline{\varphi}}$ for some n_0, i.e. $\overline{\overline{\pi}}_{n_0}^{\infty} \circ \psi = \overline{\overline{\varphi}}$. Now consider the C^*-subalgebras

$$D'_n := \varrho_n^{-1}((\overline{\overline{\pi}}_{n_0}^{n} \circ \psi)(B)) \subseteq D_n$$

and observe that the restriction of $\overline{\pi}_n^{n+1}$ to D'_n surjects onto D'_{n+1}. We also find that the direct limit $\varinjlim D'_n = \overline{\pi}_{n_0}^{\infty}(D'_{n_0})$ of this new system contains $\overline{\varphi}(A)$. Hence semiprojectivity of A allows us to lift $\overline{\varphi}$ (regarded as a map to $\varinjlim D'_n$) to D'_n for some $n \geq n_0$. Let $\sigma\colon A \to D'_n$ be a suitable partial lift, i.e. $\overline{\pi}_n^{\infty} \circ \sigma = \overline{\varphi}$, then the restriction of σ to the ideal I will be a solution to the original lifing problem for φ: The only thing we need to check is that the image of I under σ is in fact contained in C_n. But we know that $\overline{\overline{\pi}}_n^{\infty}$ is injective on $(\varrho_n \circ \sigma)(A) \subseteq (\overline{\overline{\pi}}_{n_0}^{n} \circ \psi)(B)$ since $\overline{\overline{\varphi}} = \overline{\overline{\pi}}_n^{\infty} \circ (\overline{\overline{\pi}}_{n_0}^{n} \circ \psi)$ was assumed to be injective. Hence the identity

$$(\overline{\overline{\pi}}_n^{\infty} \circ \varrho_n \circ \sigma)(i(I)) = (\varrho_\infty \circ \overline{\pi}_n^{\infty} \circ \sigma)(i(I)) = (\varrho_\infty \circ \overline{\varphi})(i(I)) = (\varrho_\infty \circ \iota_\infty)(\varphi(I)) = 0$$

confirms that $\sigma(i(I)) \subseteq i_n(C_n)$ holds. $\square$

Now assume that we are given an inductive system

$$\begin{tikzcd}
\cdots \arrow[r, two heads] & C_n \arrow[r, two heads, "\pi_n^{n+1}"] & C_{n+1} \arrow[r, two heads] & \cdots
\end{tikzcd}$$

of separable C^*-algebras with surjective connecting homomorphisms. Then each connecting map π_n^{n+1} canonically extends to a surjective *-homomorphism $\overline{\pi}_n^{n+1}$ on the level of multiplier C^*-algebras ([WO93, Theorem 2.3.9]), i.e., we automatically obtain a surjective inductive system of extensions

$$\begin{tikzcd} 0 \arrow[r] & C_n \arrow[r] \arrow[d, "\pi_n^{n+1}"'] & \mathcal{M}(C_n) \arrow[r] \arrow[d, "\overline{\pi}_n^{n+1}"'] & \mathcal{Q}(C_n) \arrow[r] \arrow[d, "\overline{\overline{\pi}}_n^{n+1}"'] & 0 . \\ 0 \arrow[r] & C_{n+1} \arrow[r] & \mathcal{M}(C_{n+1}) \arrow[r] & \mathcal{Q}(C_{n+1}) \arrow[r] & 0 \end{tikzcd}$$

We would like to apply Lemma 3.1.1 to such a system of extensions. However, the reader should be really careful when working with multipliers and direct limits at the same time since these constructions are not completely compatible: Each $\pi_n^\infty \colon C_n \to \varinjlim C_n$ extends to a *-homomorphism $\mathcal{M}(C_n) \to \mathcal{M}(\varinjlim C_n)$. The collection of these maps induces a *-homomorphism $p_{\mathcal{M}} \colon \varinjlim \mathcal{M}(C_n) \to \mathcal{M}(\varinjlim C_n)$ which is always surjective but only in trivial cases injective. The same occurs for the quotients, i.e. for the system of corona algebras $\mathcal{Q}(C_n)$. The situation can be summarized in the commutative diagram with exact rows

$$\begin{tikzcd} 0 \arrow[r] & C_n \arrow[r] \arrow[d] & \mathcal{M}(C_n) \arrow[r] \arrow[d] & \mathcal{Q}(C_n) \arrow[r] \arrow[d] & 0 \\ 0 \arrow[r] & \varinjlim C_n \arrow[r] \arrow[d, equal] & \varinjlim \mathcal{M}(C_n) \arrow[r] \arrow[d, "p_{\mathcal{M}}"] & \varinjlim \mathcal{Q}(C_n) \arrow[r] \arrow[d, "p_{\mathcal{Q}}"] & 0 \\ 0 \arrow[r] & \varinjlim C_n \arrow[r] & \mathcal{M}(\varinjlim C_n) \arrow[r] & \mathcal{Q}(\varinjlim C_n) \arrow[r] & 0 \end{tikzcd}$$

where the quotient maps $p_{\mathcal{M}}$ and $p_{\mathcal{Q}}$ are the obstacles for an application of Lemma 3.1.1. The following proposition makes these obstacles more precise.

PROPOSITION 3.1.2. *Let A and B be semiprojective C^*-algebras and*

$$0 \longrightarrow I \longrightarrow A \longrightarrow B \longrightarrow 0 \qquad [\tau]$$

a short exact sequence with Busby map $\tau \colon B \to \mathcal{Q}(I)$. Let $I \xrightarrow{\sim} \varinjlim C_n$ be an isomorphism from I to the limit of an inductive system of separable C^*-algebras C_n with surjective connecting maps. If the Busby map τ can be lifted as indicated

$$\begin{tikzcd} & \varinjlim \mathcal{Q}(C_n) \arrow[d, "p_{\mathcal{Q}}"] , \\ B \arrow[ur, dashed] \arrow[r, "\tau"] & \mathcal{Q}(I) \cong \mathcal{Q}(\varinjlim C_n) \end{tikzcd}$$

then $I \to \varinjlim C_n$ lifts to C_n for some n. If both A and B are projective, we can obtain a lift to C_1.

Proof. Keeping in mind that $p_{\mathcal{Q}}$ is the Busby map associated to the extension $0 \to \varinjlim C_n \to \varinjlim \mathcal{M}(C_n) \to \varinjlim \mathcal{Q}(C_n) \to 0$, the claim follows by combining Theorem 2.2 of [ELP99] with Lemma 3.1.1. $\square$

One special case, in which the existence of a lift for the Busby map τ as in Proposition 3.1.2 is automatic, is when the quotient B is a projective C^*-algebra. Hence we obtain a new proof for the permanence result below which has the advantage that it does not use so-called corona extendability (cf. [Lor97, Section 12.2]).

Corollary 3.1.3 ([LP98], Theorem 5.3). *Let $0 \to I \to A \to B \to 0$ be short exact. If A is (semi)projective and B is projective, then I is also (semi)projective.*

Another very specific lifting problem for which Proposition 3.1.2 applies, is the following mapping telescope contruction due to Brown.

Lemma 3.1.4. *Let a sequence $(C_k)_k$ of separable C^*-algebras be given and consider the telescope system (T_n, ϱ_n^{n+1}) associated to $\bigoplus_{k=0}^{\infty} C_k = \varinjlim_n \bigoplus_{k=0}^{n} C_k$, i.e.*

$$T_n = \left\{ f \in \mathcal{C}\left([n,\infty], \bigoplus_{k=0}^{\infty} C_k \right) : \ t \leq m \Rightarrow \ f(t) \in \bigoplus_{k=1}^{m} C_k \right\}$$

with $\varrho_n^{n+1} \colon T_n \to T_{n+1}$ the (surjective) restriction maps, so that $\varinjlim_n (T_n, \varrho_n^{n+1}) \cong \bigoplus_{k=1}^{\infty} C_k$. Then both canonical quotient maps in the diagram

$$\begin{tikzcd}
0 \arrow[r] & \varinjlim T_n \arrow[r] \arrow[d, equal] & \varinjlim \mathcal{M}(T_n) \arrow[r] \arrow[d, "p_{\mathcal{M}}"] & \varinjlim \mathcal{Q}(T_n) \arrow[r] \arrow[d, "p_{\mathcal{Q}}"] & 0 \\
0 \arrow[r] & \varinjlim T_n \arrow[r] & \mathcal{M}(\varinjlim T_n) \arrow[r] \arrow[u, dotted, bend left] & \mathcal{Q}(\varinjlim T_n) \arrow[r] \arrow[u, dotted, bend left] & 0
\end{tikzcd}$$

split.

Proof. It suffices to produce a split for $p_{\mathcal{M}}$ which is the identity on $\varinjlim T_n$. Under the identification $\varinjlim T_n \cong \bigoplus_{k=0}^{\infty} C_k$ we have $\mathcal{M}(\varinjlim T_n) \cong \prod_{k=0}^{\infty} \mathcal{M}(C_k)$. One checks that

$$T_n = \bigoplus_{k=0}^{n} \mathcal{C}([n,\infty], C_k) \oplus \bigoplus_{k>n} \mathcal{C}_0((k,\infty], C_k)$$

and hence

$$\prod_{k=0}^{\infty} \mathcal{C}([\max\{n,k\},\infty], \mathcal{M}(C_k)) \subset \mathcal{M}(T_n).$$

It follows that the sum of embeddings as constant functions

$$\prod_{k=0}^{\infty} \mathcal{M}(C_k) \to \prod_{k=0}^{\infty} \mathcal{C}([\max\{n,k\},\infty], \mathcal{M}(C_k)) \subset \mathcal{M}(T_n)$$

defines a split for the quotient map $\varinjlim \mathcal{M}(T_n) \to \mathcal{M}(\varinjlim T_n)$. It is easily verified that this split is the identity on $\bigoplus_{k=1}^{\infty} C_k$. $\square$

Remark 3.1.5 (Lifting the Busby map). *Given an extension $0 \to I \to A \to B \to 0$ with both A and B semiprojective, the associated Busby map does in general not lift as in 3.1.2. However, there are a number of interesting situations where it does lift and we therefore can use Propostion 3.1.2 to obtain lifting properties for the ideal I. One such example is studied in [End14], where it is (implicitly) shown that the Busby map lifts if B is a finite-dimensional C^*-algebra. This observation leads to the fact that semiprojectivity passes to ideals of finite codimension. Further examples will be given in section 4, where we study Busby maps associated to extensions by homogeneous ideals and identify conditions which guarantee that 3.1.2 applies.*

3.2 Direct limits which preserve semiprojectivity

3.2.1 Weakly conditionally projective homomorphisms

The following definition characterizes *-homomorphisms along which lifting solutions can be extended in an approximate manner. This type of maps is implicitly used in [CD10] and [ST12] in the special case of finitely presented, commutative C^*-algebras.

Definition 3.2.1. *A *-homomorphism $\varphi\colon A \to B$ is weakly conditionally projective if the following holds: Given $\epsilon > 0$, a finite subset $F \subset A$ and a commuting square*

$$\begin{tikzcd} A \arrow[r, "\psi"] \arrow[d, "\varphi"'] & D \arrow[d, "\pi"] \\ B \arrow[r, "\varrho"] & D/J, \end{tikzcd}$$

there exists a *-homomorphism $\psi'\colon B \to D$ as indicated

$$\begin{tikzcd} A \arrow[r, "\psi"] \arrow[d, "\varphi"'] & D \arrow[d, "\pi"] \\ B \arrow[r, "\varrho"] \arrow[ur, dashed, "\psi'"] & D/J \end{tikzcd}$$

which satisfies $\pi \circ \psi' = \varrho$ and $\|(\psi' \circ \varphi)(a) - \psi(a)\| < \epsilon$ for all $a \in F$.

The definition above is a weakening of the notion of conditionally projective morphisms, as introduced in section 5.3 of [ELP98], where one asks the homomorphism ψ' in 3.2.1 to make both triangles of the lower diagram to commute exactly. While conditionally projective morphisms are extremely rare (even when working with projective C^*-algebras, cf. the example below), there is a sufficient supply of weakly conditionally projective ones, as we will show in the next section.

Example 3.2.2. *The inclusion map* $\mathrm{id}\oplus 0\colon \mathcal{C}_0(0,1] \to \mathcal{C}_0(0,1]\oplus\mathcal{C}_0(0,1]$ *is weakly conditionally projective but not conditionally projective. This can be illustrated by considering the commuting square*

$$\begin{tikzcd}
\mathcal{C}_0(0,1] \arrow[rr, "\psi"] \arrow[d, "\mathrm{id}\oplus 0"'] & & \mathcal{C}_0(0,3) \arrow[d, "\pi"] \\
\mathcal{C}_0(0,1]\oplus\mathcal{C}_0[2,3) \arrow[rr, equal] & & \mathcal{C}_0(0,1]\oplus\mathcal{C}_0[2,3)
\end{tikzcd}$$

where π *is the restriction map and* ψ *is given by sending the canonical generator* t *of* $\mathcal{C}_0(0,1]$ *to the function*

$$(\psi(t))(s) = \begin{cases} s & \text{if } s\leq 1 \\ 1-s & \text{if } 1<s\leq 2 \\ 0 & \text{if } 2\leq s \end{cases}.$$

It is clear that there is no lift for the generator of $\mathcal{C}_0[2,3)$ *which is orthogonal to* $\psi(t)$*. This shows that the map* $\mathrm{id}\oplus 0$ *is not conditionally projective. However, after replacing* $\psi(t)$ *with* $(\psi(t)-\epsilon)_+$ *for any* $\epsilon>0$*, finding an orthogonal lift for the generator of the second summand is no longer a problem. Using this idea, it will be shown in Proposition 3.2.4 that* $\mathrm{id}\oplus 0$ *is in fact weakly conditionally projective.*

If A is a (semi)projective C^*-algebra and $\varphi\colon A\to B$ is weakly conditionally projective, then B is of course also (semi)projective. The next lemma shows that (semi)projectivity is even preserved along a sequence of such maps. Its proof is of an approximate nature and relies on a one-sided approximate intertwining argument (cf. section 2.3 of [Rør02]), a technique borrowed from the Elliott classification program.

Lemma 3.2.3. *Suppose* $A_1 \xrightarrow{\varphi_1^2} A_2 \xrightarrow{\varphi_2^3} A_3 \xrightarrow{\varphi_3^4} \cdots$ *is an inductive system of separable* C^**-algebras. If* A_1 *is (semi)projective and all connecting maps* φ_n^{n+1} *are weakly conditionally projective, then the limit* $A_\infty = \varinjlim(A_n,\varphi_n^{n+1})$ *is also (semi)projective.*

Proof. We will only consider the projective case, the statement for the semiprojective case is proven analogously with obvious modifications. Choose finite subsets $F_n\subset A_n$ with $\varphi_n^{n+1}(F_n)\subseteq F_{n+1}$ such that the union $\bigcup_{m=n}^\infty(\varphi_n^m)^{-1}(F_m)$ is dense in A_n for all n. Further let $(\epsilon_n)_n$ be a sequence in $\mathbb{R}_{>0}$ with $\sum_{n=1}^\infty \epsilon_n<\infty$. Now let $\varrho\colon A_\infty\to D/J$ be a *-homomorphism to some quotient C^*-algebra D/J. By projectivity of A_1 there is a *-homomorphism $s_1\colon A_1\to D$ with $\pi\circ s_1=\varrho\circ\varphi_1^\infty$. Since the maps φ_n^{n+1} are weakly conditionally projective, we can inductively choose $s_{n+1}\colon A_{n+1}\to D$ with $\pi\circ s_{n+1}=\varrho\circ\varphi_{n+1}^\infty$ such that

$$\|s_n(a)-(s_{n+1}\circ\varphi_n^{n+1})(a)\|<\epsilon_n$$

holds for all $a \in F_n$. It is now a standard computation (and therefore ommited) to check that $((s_m \circ \varphi_n^m)(x))_m$ is a Cauchy sequence in D for every $x \in F_n$. Furthermore, the induced map $\varphi_n^\infty(x) \mapsto \lim_m (s_m \circ \varphi_n^m)(x)$ extends from the dense subset $\bigcup_n \varphi_n^\infty(F_n)$ to a *-homomorphism $s\colon A_\infty \to D$.

$$\begin{tikzcd}
A_n \arrow[r, "s_n"] \arrow[d, "\varphi_n^{n+1}"'] & D \arrow[dd, two heads, "\pi"] \\
A_{n+1} \arrow[ur, dashed, "s_{n+1}"] \arrow[d, "\varphi_{n+1}^\infty"'] & \\
A_\infty \arrow[uur, dotted, "s"] \arrow[r, "\varrho"] & D/J
\end{tikzcd}$$

Since each s_n lifts π, the same holds for their pointwise limit, i.e. the limit map s satisfies $\pi \circ s = \varrho$. This shows that A_∞ is projective. $\square$

3.2.2 Adding non-commutative edges

In order to make Lemma 3.2.3 a useful tool for constructing semiprojective C^*-algebras, we have to ensure the existence of weakly projective *-homomorphisms as defined in 3.2.1. The examples we work out in this section arise in special pullback situations where one 'adds a non-commutative edge' to a given C^*-algebra A. By this we mean that we form the pullback of A and $\mathcal{C}([0,1]) \otimes \mathbb{M}_n$ over a n-dimensional representation of A and the evaluation map ev_0. In the special case of $A = \mathcal{C}(X)$ being a commutative C^*-algebra and $n = 1$ this pullback construction already appeared in [CD10] and [ST12] where it indeed corresponds to attaching an egde $[0,1]$ at one point to the space X. Here we show that the map obtained by extending elements of A as constant functions onto the attached non-commutative edge gives an example of a weakly conditionally projective *-homomorphism. As an application, we observe that the AF-telescopes studied in [LP98] arise from weakly projective *-homomorphisms and hence projectivity of these algebras is a direct consequence of Lemma 3.2.3.

Adapting notation from [ELP98], we set

$$T(\mathbb{C}, G) = \{f \in \mathcal{C}_0((0,2], G) : \quad t \leq 1 \Rightarrow f(t) \in \mathbb{C} \cdot 1_G\},$$

$$S(\mathbb{C}, G) = \{f \in \mathcal{C}_0((0,2), G) : \quad t \leq 1 \Rightarrow f(t) \in \mathbb{C} \cdot 1_G\}$$

for G a unital C^*-algebra. We further write

$$T(\mathbb{C}, G, F) = \left\{ f \in \mathcal{C}_0((0,3], F) : \begin{array}{l} t \leq 2 \Rightarrow f(t) \in G \\ t \leq 1 \Rightarrow f(t) \in \mathbb{C} \cdot 1_G \end{array} \right\}$$

with respect to a fixed inclusion $G \subseteq F$. We have the diagram

$$\begin{tikzcd} T(\mathbb{C},G,F) \arrow[r] \arrow[d] & \mathcal{C}([2,3],F) \arrow[d, "\mathrm{ev}_2"] \\ T(\mathbb{C},G) \arrow[r, "\mathrm{ev}_2"] & F \end{tikzcd}$$

which is a special case of the pullback situation considered in the next proposition. However, this example is in some sense generic and implementing it into the general situation is an essential part of proving the following.

PROPOSITION 3.2.4. *Given a (semi)projective C^*-algebra Q and a *-homomorphism $\tau\colon Q \to \mathbb{M}_n$, the following holds:*

1. *The pullback P over τ and $\mathrm{ev}_0\colon \mathcal{C}([0,1],\mathbb{M}_n) \to \mathbb{M}_n$, i.e.*

$$P = \{(q,f) \in Q \oplus \mathcal{C}([0,1],\mathbb{M}_n) :\ \tau(q) = f(0)\},$$

   is (semi)projective.

2. *The canonical split $s\colon Q \to P$, $q \mapsto (q, \tau(q) \otimes 1_{[0,1]})$ is weakly conditionally projective.*

Proof. (1) Semiprojectivity of the pullback P follows from [End14, Corollary 3.4]. Since P is homotopy equivalent to Q, the projective statement follows from the semiprojective one using [Bla12, Corollary 5.2].

(2) For technical reasons we identify the attached interval $[0,1]$ with $[2,3]$ and consider the pullback

$$\begin{tikzcd} P \arrow[r, dashed] \arrow[d, dashed] & \mathcal{C}([2,3],\mathbb{M}_n) \arrow[d, "\mathrm{ev}_2"] \\ Q \arrow[r, "\tau"] & \mathbb{M}_n \end{tikzcd}$$

with $s\colon Q \to P$, $q \mapsto (q, \tau(q) \otimes 1_{[2,3]})$ instead. Denote by $G \subseteq \mathbb{M}_n$ the image of τ. According to [ELP98, Theorem 2.3.3], we can find a *-homomorphism $\overline{\varphi}\colon T(\mathbb{C},G) \to Q$ such that

$$\begin{tikzcd} 0 \arrow[r] & \ker(\tau) \arrow[r] & Q \arrow[r, "\tau"] & G \arrow[r] \arrow[d, equal] & 0 \\ 0 \arrow[r] & S(\mathbb{C},G) \arrow[r] \arrow[u] & T(\mathbb{C},G) \arrow[r, "ev_2"] \arrow[u, "\overline{\varphi}"] & G \arrow[r] & 0 \end{tikzcd}$$

commutes and $\overline{\varphi}_{|S(\mathbb{C},G)}$ is a proper *-homomorphism to $\ker(\tau)$ (meaning that the hereditary subalgebra generated by its image is all of $\ker(\tau)$). Using the pullback property of P, $\overline{\varphi}$ can be extended to $\varphi\colon T(\mathbb{C},G,\mathbb{M}_n)\to P$ such that

$$\begin{tikzcd}
0 \arrow[r] & \mathcal{C}_0((2,3],\mathbb{M}_n) \arrow[r] \arrow[d, equal] & P \arrow[r, shift right] & Q \arrow[l, bend right, "s"'] \arrow[r] & 0 \\
0 \arrow[r] & \mathcal{C}_0((2,3],\mathbb{M}_n) \arrow[r] & T(\mathbb{C},G,\mathbb{M}_n) \arrow[u, "\varphi"] \arrow[r, shift right] & T(\mathbb{C},G) \arrow[u, "\overline{\varphi}"'] \arrow[l, shift right, "s'"'] \arrow[r] & 0
\end{tikzcd}$$

commutes. In particular we have $\varphi\circ s'=s\circ\overline{\varphi}$, where s' is the canonical split which simply extends functions constantly onto $[2,3]$.
Choose generators $f_1,...,f_l$ of norm 1 for $\mathcal{C}_0((2,3],\mathbb{M}_n)$ and generators $g_1,...,g_k$ of norm 1 for $T(\mathbb{C},G)$. We need the following 'softened' versions of P: For $\delta>0$ we consider the universal C^*-algebra

$$P_\delta = C^*\left(\left\{f^\delta, q^\delta : f\in\mathcal{C}_0((2,3],\mathbb{M}_n), q\in Q\right\} | \mathcal{R}_{\mathcal{C}_0((2,3],\mathbb{M}_n)} \& \mathcal{R}_Q \& \mathcal{R}_\delta\right)$$

which is generated by copies of $\mathcal{C}_0((2,3],\mathbb{M}_n)$ and Q (here $\mathcal{R}_{\mathcal{C}_0((2,3],\mathbb{M}_n)}, \mathcal{R}_Q$ denote all the relations from $\mathcal{C}_0((2,3],\mathbb{M}_n)$ resp. from Q) and additional, finitely many relations

$$\mathcal{R}_\delta = \left\{\|f_i^\delta(\overline{\varphi}(g_j))^\delta - (f_i(g_j(2)\otimes 1_{[2,3]}))^\delta\| \le \delta\right\}_{\substack{1\le i\le l\\ 1\le j\le k}}.$$

Note that $P=\varinjlim P_\delta$ with respect to the canonical surjections $p_{\delta,\delta'}\colon P_\delta\to P_{\delta'}$ (for $\delta>\delta'$) and denote the induced maps $P_\delta\to P, f^\delta\mapsto f, q^\delta\mapsto s(q)$ by $p_{\delta,0}$. Since P is semiprojective by part (1) of this proposition, we can find a partial lift $j_\delta\colon P\to P_\delta$ for some $\delta>0$, i.e. $p_{\delta,0}\circ j_\delta=\mathrm{id}_P$.
Now let a finite set $F=\{x_1,...,x_m\}\subseteq Q$ and $\epsilon>0$ and be given. Denoting the inclusions $Q\to P_\delta, q\mapsto q^\delta$ by s_δ, we can (after decreasing δ if necessary) assume that $\|s_\delta(x_i)-(j_\delta\circ s)(x_i)\|\le\epsilon$ holds for all $1\le i\le m$. Now given any commuting square

$$\begin{tikzcd}
Q \arrow[r, "\psi"] \arrow[d, "s"] & D \arrow[d, "\pi"] \\
P \arrow[r, "\varrho"] & D/J
\end{tikzcd}$$

it only remains to construct a *-homomorphism $\psi_\delta\colon P_\delta\to D$ such that in the diagram

$$\begin{tikzcd}
Q \arrow[rr, "\psi"] \arrow[dr, "s_\delta"] \arrow[dd, "s"'] & & D \arrow[dd, "\pi"] \\
& P_\delta \arrow[ur, dashed, "\psi_\delta"] \arrow[dl, "p_{\delta,0}"'] & \\
P \arrow[ur, bend right, "j_\delta"'] \arrow[rr, "\varrho"'] & & D/J
\end{tikzcd}$$

the upper central triangle and the lower right triangle commute.
We consider the following subalgebras of $T(\mathbb{C},G)$ and $S(\mathbb{C},G)$ for any $\eta>0$:

$$\begin{aligned} T_\eta(\mathbb{C},G) &= \{f\in T(\mathbb{C},G): f \text{ is constant on } (0,\eta]\cup[2-\eta,2]\} \\ S_\eta(\mathbb{C},G) &= \{f\in S(\mathbb{C},G): f \text{ is constant (=0) on } (0,\eta]\cup[2-\eta,2]\} \end{aligned}$$

Since

$$T(\mathbb{C},G)=\overline{\bigcup_{\eta>0} T_\eta(\mathbb{C},G)}$$

we find $0<\eta<\frac{1}{2}$ and elements $\tilde{g}_j\in T_\eta(\mathbb{C},G)$ with $\tilde{g}_j(2)=g_j(2)$ and $\|g_j-\tilde{g}_j\|<\delta$ for every $1\le j\le k$. Let $h\in T(\mathbb{C},G)$ be the scalar-valued function which equals 1_G on $[\eta,2-\eta]$, satisfies $h(0)=h(2)=0$ and is linear in between. Consider the hereditary C^*-subalgebra $D'=\overline{(1-(\psi\circ\overline{\varphi})(h))D(1-(\psi\circ\overline{\varphi})(h))}$ and define

$$D'':=(\psi\circ\overline{\varphi})(T_\eta(\mathbb{C},G))+D'\subseteq D.$$

Then $(\psi\circ\overline{\varphi})(S_\eta(\mathbb{C},G))$ and D' are orthogonal ideals in D'' because h is central in $T(\mathbb{C},G)$. We further have $(\varrho\circ\varphi)(\mathcal{C}_0((2,3],\mathbb{M}_n))\subseteq\pi(D')$ and hence obtain a commutative diagram

$$\begin{tikzcd}
0 \arrow[r] & (\psi\circ\overline{\varphi})(S_\eta(\mathbb{C},G)) \arrow[r] \arrow[d, two heads, "\pi"] & D'' \arrow[r] \arrow[d, two heads, "\pi"] & H_D+D' \arrow[r] \arrow[d, two heads, dashed] & 0 \\
0 \arrow[r, tail] & (\varrho\circ\varphi\circ s')(S_\eta(\mathbb{C},G)) \arrow[r] & \pi(D'') \arrow[r] & H_{D/J}+\pi(D') \arrow[r, tail] & 0 \\
0 \arrow[r] & s'(S_\eta(\mathbb{C},G)) \arrow[r] \arrow[u, "\varrho\circ\varphi"] & s'(T_\eta(\mathbb{C},G))+\mathcal{C}_0((2,3],\mathbb{M}_n) \arrow[r] \arrow[u, "\varrho\circ\varphi"] & \hat{T}(G,\mathbb{M}_n) \arrow[r] \arrow[u, dashed] & 0
\end{tikzcd}$$

where H_D and $H_{D/J}$ are finite-dimensional C^*-algebras given by

$$H_D=(\psi\circ\overline{\varphi})(T_\eta(\mathbb{C},G))/(\psi\circ\overline{\varphi})(S_\eta(\mathbb{C},G)),$$

$$H_{D/J}=(\varrho\circ\varphi\circ s')(T_\eta(\mathbb{C},G))/(\varrho\circ\varphi\circ s')(S_\eta(\mathbb{C},G))$$

and $\hat{T}(G,\mathbb{M}_n)$ denotes what is called a crushed telescope in [ELP98]:

$$\hat{T}(G,\mathbb{M}_n)=\{f\in\mathcal{C}([2,3],\mathbb{M}_n)\colon f(2)\in G\}$$

By [ELP98, Proposition 6.1.1], the embedding $G\to\hat{T}(G,\mathbb{M}_n)$ as constant functions is a conditionally projective map (in the sense of [ELP98, Section 5.3]). It is hence possible to extend the map $G\xrightarrow{\sim}T_\eta(\mathbb{C},G)/S_\eta(\mathbb{C},G)\longrightarrow H_D\subset H_D+D'$ to a *-homomorphism

$\psi' \colon \hat{T}(G, \mathbb{M}_n) \to H_D + D'$ such that the diagram with exact rows

$$\begin{tikzcd}
0 \arrow[r] & D' \arrow[r] \arrow[d, "\pi"'] & H_D + D' \arrow[r, shift left] \arrow[d] & H_D \arrow[r] \arrow[d] \arrow[l, shift left] & 0 \\
0 \arrow[r] & \pi(D') \arrow[r] & H_{D/J} + \pi(D') \arrow[r] & H_{D/J} \arrow[r] & 0 \\
0 \arrow[r] & \mathcal{C}_0((2,3], \mathbb{M}_n) \arrow[r] \arrow[u, "\varrho\circ\varphi"] & \hat{T}(G, \mathbb{M}_n) \arrow[r, shift left] \arrow[u] \arrow[uu, dashed, bend left, "\psi'"] & G \arrow[r] \arrow[u] \arrow[l, shift left, "\mathrm{ev}_2"] \arrow[uu, bend left] & 0
\end{tikzcd}$$

commutes. In particular, ψ' restricts to a *-homomorphism $\mathcal{C}_0((2,3], \mathbb{M}_n) \to D'$ which we will also denote by ψ'. But then a diagram chase confirms that

$$\psi'(f_i) \cdot (\psi \circ \overline{\varphi})(\tilde{g}_j) = \psi'(f_i \cdot (\tilde{g}_j(2) \otimes 1_{[2,3]}))$$

holds for every i, j. Finally, define $\psi_\delta \colon P_\delta \to D$ by

$$q^\delta \mapsto \psi(q) \quad \text{and} \quad f_i^\delta \mapsto \psi'(f_i).$$

It needs to be checked that ψ_δ is well-defined, i.e. that the elements $\psi_\delta(f_i^\delta)$ and $\psi_\delta(\overline{\varphi}(g_j)^\delta)$ satisfy the relations $\mathcal{R}_\delta$:

$$\begin{aligned}
& \|\psi_\delta(f_i^\delta)\psi_\delta(\overline{\varphi}(g_j)^\delta) - \psi_\delta((f_i(g_j(2) \otimes 1_{[2,3]}))^\delta)\| \\
= \; & \|\psi'(f_i)((\psi \circ \overline{\varphi})(g_j)) - \psi'(f_i(g_j(2) \otimes 1_{[2,3]}))\| \\
\leq \; & \|\psi'(f_i)(\psi \circ \overline{\varphi})(\tilde{g}_j) - \psi'(f_i(\tilde{g}_j(2) \otimes 1_{[2,3]}))\| + \|f_i\| \cdot \|g_j - \tilde{g}_j\| < \delta
\end{aligned}$$

Since we also have $\psi_\delta \circ s_\delta = \psi$ and $\pi \circ \psi_\delta = \varrho \circ p_{\delta,0}$, the proof is hereby complete. ☐

One example, where pullbacks as in 3.2.4 show up, is the class of so-called AF-telescopes defined by Loring and Pedersen:

Definition 3.2.5 ([LP98]). *Let $A = \overline{\bigcup A_n}$ be the inductive limit of an increasing union of finite-dimensional C^*-algebras A_n. We define the AF-telescope associated to this AF-system as*

$$T(A) = \{f \in \mathcal{C}_0((0,\infty], A) : \; t \leq n \Rightarrow f(t) \in A_n\}.$$

We have an obvious limit structure for $T(A) = \varinjlim T(A_k)$ over the finite telescopes

$$T(A_k) = \{f \in \mathcal{C}_0((0,k], A_k)) : \; t \leq n \Rightarrow f(t) \in A_n\}.$$

Now the embedding of $T(A_k)$ into $T(A_{k+1})$ is given by extending the elements of $T(A_k)$ constantly onto the attached interval $[k, k+1]$. This is nothing but a finite composition of maps as in part (2) of 3.2.4. Hence the connecting maps in the system of finite telescopes are weakly conditionally projective and using Lemma 3.2.3 we recover [LP98, Theorem 7.2]:

Corollary 3.2.6. *All AF-telescopes are projective.*

In contrast to the original proof we didn't have to work out any description of the telescopes by generators and relations. Such a description would have to encode the structure of each A_n as well as the inlusions $A_n \subset A_{n+1}$ (i.e., the Bratteli-diagram of the system). Showing that such an infinite set of generators and relations gives rise to a projective C^*-algebra is possible but complicated. Instead we showed that these algebras are build up from the projective C^*-algebra $T(A_0)$=0 using operations which preserve projectivity.

4 Extensions by homogeneous C^*-algebras

In this section we study extensions by (trivially) homogeneous C^*-algebras, i.e. extensions of the form

$$0 \longrightarrow \mathcal{C}_0(X, \mathbb{M}_N) \longrightarrow A \longrightarrow B \longrightarrow 0.$$

Our final goal is to understand the behavior of semiprojectivity along such extensions, and we will eventually achieve this in Theorem 4.3.2.

4.1 Associated retract maps

Identifying X with an open subset of $\mathrm{Prim}(A)$, we make the following definition of an associated retract map. This map will play a key role in our study of extensions.

Definition 4.1.1. *Let X be locally compact space with connected components $(X_i)_{i \in I}$ and*

$$0 \longrightarrow \mathcal{C}_0(X, \mathbb{M}_N) \longrightarrow A \longrightarrow B \longrightarrow 0$$

a short exact sequence of C^*-algebras. We define the (set-valued) retract map R associated to the extension to be the map

$$R\colon \mathrm{Prim}(A) \to 2^{\mathrm{Prim}(B)}$$

given by

$$R(z) = \begin{cases} z & \text{if } z \in \mathrm{Prim}(B), \\ \partial X_i = \overline{X_i} \backslash X_i & \text{if } z \in X_i \subseteq X. \end{cases}$$

Note that R defined as above takes indeed values in $2^{\mathrm{Prim}(B)}$ because the connected components X_i are always closed in X. However, in our cases of interest the components X_i will actually be clopen in X (e.g. if X is locally connected) so that we have a topological decomposition $X = \bigsqcup_i X_i$.

4.1.1 REGULARITY PROPERTIES FOR SET-VALUED MAPS

Let X, Y be sets and $S\colon X \to 2^Y$ a set-valued map. We say that S has pointwise finite image if $S(x) \subseteq Y$ is a finite set for every $x \in X$. If furthermore X and Y are topological spaces, we will use the following notion of semicontinuity for S (cf. [AF90, Section 1.4]).

DEFINITION + LEMMA 4.1.2. *Let X, Y be topological spaces. A set-valued map $S\colon X \to 2^Y$ is lower semicontinuous if one of the following equivalent conditions holds:*

(i) $\{x \in X \colon S(x) \subseteq B\}$ *is closed in X for every closed $B \subseteq Y$.*

(ii) For every neighborhood $N(\overline{y})$ of $\overline{y} \in S(\overline{x})$ there exists a neighborhood $N(\overline{x})$ of $\overline{x}$ with $S(x) \cap N(\overline{y}) \neq \emptyset$ for every $x \in N(\overline{x})$.

(iii) For every net $(x_\lambda)_{\lambda\in\Lambda} \subset X$ with $x_\lambda \to x$ and every $y \in S(x)$ there exists a net $(y_\mu)_{\mu\in M} \subset \{S(x_\lambda)\colon \lambda \in \Lambda\}$ such that $y_\mu \to y$.

Proof. $(i) \Rightarrow (ii)$: Let $N(\overline{y})$ be an open neighborhood of $\overline{y} \in S(\overline{x})$. Then $\{x \in X : S(x) \subset Y \backslash N(\overline{y})\}$ is closed and does not contain $\overline{x}$. Hence we find an open neighborhood $N(\overline{x})$ of $\overline{x}$ in $X\backslash\{x \in X : S(x) \subset Y\backslash N(\overline{y})\} = \{x \in X : S(x) \cap N(\overline{y}) \neq \emptyset\}$.
$(ii) \Rightarrow (iii)$: Denote by $\mathcal{N}$ the family of neighborhoods of y ordered by reversed inclusion. Set $M = \{(\lambda, N) \in \Lambda \times \mathcal{N}\colon S(x_{\lambda'}) \cap N \neq \emptyset\ \forall\ \lambda' \geq \lambda\}$, then by assumption M is nonempty and directed with respect to the partial order $(\lambda_1, N_1) \leq (\lambda_2, N_2)$ iff $\lambda_1 \leq \lambda_2$ and $N_2 \subseteq N_1$. Now pick a $y_{(\lambda,N)} \in S(x_\lambda) \cap N$ for each $(\lambda, N) \in M$, then $(y_\mu)_{\mu\in M}$ constitutes a suitable net converging to y.
$(iii) \Rightarrow (i)$: Let a closed set $B \subseteq Y$ and $(x_\lambda)_{\lambda\in\Lambda} \subset \{x \in X : S(x) \subseteq B\}$ with $x_\lambda \to \overline{x}$ be given. Then for any $\overline{y} \in S(\overline{x})$ we find a net $y_\mu \to \overline{y}$ with $(y_\mu) \subset \{S(x_\lambda) : \lambda \in \Lambda\} \subset B$. Since B is closed we have $\overline{y} \in B$ showing that $S(\overline{x}) \subset B$. ☐

REMARK 4.1.3. *An ordinary (i.e. a single-valued) map is evidently lower semicontinuous in the sense above if and only if it is continuous. If both spaces X and Y are first countable, we may use sequences instead of nets in condition (iii).*

Examples of set-valued maps that are lower semicontinuous in the sense above arise from split extensions by homogeneous C^*-algebras as follows.

EXAMPLE 4.1.4. *Let a split-exact sequence of separable C^*-algebras*

$$\begin{tikzcd} 0 \arrow[r] & \mathcal{C}_0(X, \mathbb{M}_n) \arrow[r] & A \arrow[r, "\pi"'] & B \arrow[r] \arrow[l, bend right, "s"'] & 0 \end{tikzcd}$$

be given and consider the set-valued map $R_s\colon \mathrm{Prim}(A) \to 2^{\mathrm{Prim}(B)}$ given by

$$R_s(z) = \begin{cases} z & \text{if } z \in \mathrm{Prim}(B) \\ \{[\pi_{z,1}], ..., [\pi_{z,r(z)}]\} & \text{if } z \in X \end{cases}$$

where $\pi_{z,1} \oplus ... \oplus \pi_{z,r(z)}$ *is the decomposition of* $B \xrightarrow{s} A \to \mathcal{C}_b(X, \mathbb{M}_n) \xrightarrow{\mathrm{ev}_z} \mathbb{M}_n$ *into irreducible summands. Then* R_s *is lower semicontinuous in the sense of 4.1.2.*

Proof. We verify condition (ii) of 4.1.2: Let $z_n \to \overline{z}$ in $\mathrm{Prim}(A)$ and a neighborhood $N(\overline{y})$ of $\overline{y} \in R_s(\overline{z})$ in $\mathrm{Prim}(B)$ be given. By Lemma 2.2.4 we may assume that $N(\overline{y})$ is of the form $\{z \in \mathrm{Prim}(B)\colon \check{b}(z) > 1/2\}$ for some $b \in B$. By definition of R_s, we find $\check{s(b)}(z) = \max_{y \in R_s(z)} \check{b}(y)$ for all $z \in \mathrm{Prim}(A)$. Hence $N(\overline{z}) = \{z \in \mathrm{Prim}(A)\colon \check{s(b)} > 1/2\}$ constitutes a neighborhood of $\overline{z}$ in $\mathrm{Prim}(A)$ which satisfies 4.1.2 (ii). □

Note that the retract map R_s in 4.1.4 highly depends on the choice of splitting s while the retract map R from 4.1.1 is associated to the underlying extension in a natural way. It is the goal of section 4.1.2 to find a splitting s such that $R = R_s$ holds. This is, however, not always possible. It can even happen that the underlying extension splits while R is not of the form R_s for any splitting s (cf. remark 4.3.3). Under suitable conditions, we will at least be able to arrange $R = R_s$ outside of a compact set $K \subset X$, i.e. we can find a (not necessarily multiplicative) splitting map s such that $B \xrightarrow{s} A \to \mathcal{C}_b(X, \mathbb{M}_n)$ is multiplicative on $X \backslash K$ so that $R_s(x)$ is still well-defined and coincides with $R(x)$ for all $x \in \mathrm{Prim}(A) \backslash K$.

4.1.2 Lifting the Busby map

In this section we identify conditions on an extension

$$\begin{tikzcd} 0 \arrow[r] & \mathcal{C}_0(X, \mathbb{M}_N) \arrow[r] & A \arrow[r] & B \arrow[r] \arrow[l, bend right, dotted, "?"'] & 0 \end{tikzcd} \qquad [\tau]$$

which allow us to contruct a splitting $s\colon B \to A$. This is evidently the same as asking for a lift of the corresponding Busby map τ as indicated on the left of the commutative diagram

$$\begin{tikzcd} & \mathcal{C}(\beta X, \mathbb{M}_N) \arrow[r, equal] \arrow[d, "\varrho"] & \prod_{i \in I} \mathcal{C}(\beta X_i, \mathbb{M}_N) \arrow[d] \\ B \arrow[ur, dotted, "s"] \arrow[r, "\tau"] \arrow[rr, bend right, "\oplus \tau_i"'] & \mathcal{C}(\chi(X), \mathbb{M}_N) \arrow[r, two heads] & \prod_{i \in I} \mathcal{C}(\chi(X_i), \mathbb{M}_N). \end{tikzcd}$$

We will produce a suitable lift of τ in two steps:

1. For every component X_i of X, we trivialize the map $\tau_i\colon B \to \mathcal{C}(\chi(X_i), \mathbb{M}_N)$, i.e. we conjugate it to a constant map, so that it can be lifted to $\mathcal{C}(\beta X_i, \mathbb{M}_N)$. This step requires the associated retract map R from 4.1.1 to have pointwise finite image and the spaces $\chi(X_i)$ to be connected and low-dimensional.

2. We extend the collection of lifts for the τ_i's to a lift for τ. Here we need the associated retract map R to be lower semicontinuous.

In many cases of interest, the spaces $\chi(X_i)$ will not be connected, so that we have to modify the first step of the lifting process. This results in the fact that we cannot find a (multiplicative) split s in general. Instead we will settle for a lift s of τ with slightly weaker multiplicative properties.
First we give the connection between the retract map R and the Busby map τ of the extension.

LEMMA 4.1.5. *Let a short exact sequence*

$$0 \longrightarrow I \longrightarrow A \xrightarrow{\pi} B \longrightarrow 0$$

with Busby map $\tau\colon B \to \mathcal{Q}(I)$ *be given. Identifying* $\operatorname{Prim}(I)$ *with the open subset* $\{J | I \nsubseteq J\}$ *of* $\operatorname{Prim}(A)$ *and denoting by* $\partial \operatorname{Prim}(I)$ *its boundary in* $\operatorname{Prim}(A)$, *the following statements hold:*

(i) $J \in \partial \operatorname{Prim}(I) \Leftrightarrow I + I^\perp \subseteq J$ *for every* $J \in \operatorname{Prim}(A)$,

(ii) $\partial \operatorname{Prim}(I) = \{J\colon \ker(\tau \circ \pi) \subseteq J\} \cong \operatorname{Prim}(\tau(B))$.

If in addition I is subhomogeneous, we further have

(iii) $|\partial \operatorname{Prim}(I)| < \infty \Leftrightarrow \dim(\tau(B)) < \infty$.

Proof. For (i) it suffices to check that $\operatorname{Prim}(I^\perp) = \operatorname{Prim}(A) \backslash \overline{\operatorname{Prim}(I)}$ where $I^\perp$ denotes the annihilator of I in A. But this follows directly from the definition of the Jacobson topology on $\operatorname{Prim}(A)$:

$$\begin{aligned} J \notin \overline{\operatorname{Prim}(I)} &\Leftrightarrow \bigcap_{K \in \operatorname{Prim}(I)} K \nsubseteq J \\ &\Leftrightarrow \exists x \in A\colon x \notin J \text{ while } \|x\|_K = 0 \;\forall\, K \in \operatorname{Prim}(I) \\ &\Leftrightarrow \exists x \in I^\perp\colon x \notin J \\ &\Leftrightarrow I^\perp \nsubseteq J \\ &\Leftrightarrow J \in \operatorname{Prim}(I^\perp). \end{aligned}$$

As $I^\perp = \ker(A \to \mathcal{M}(I))$, one finds $\ker(\tau \circ \pi) = I + I^\perp$. Together with (i) this shows

$$\begin{aligned} \operatorname{Prim}(\tau(B)) &= \operatorname{Prim}((\tau \circ \pi)(A)) \\ &\cong \{J \in \operatorname{Prim}(A)\colon \ker(\tau \circ \pi) \subseteq J\} \\ &= \{J \in \operatorname{Prim}(A)\colon I + I^\perp \subseteq J\} \\ &= \partial \operatorname{Prim}(I). \end{aligned}$$

For the last statement note that if all irreducible representations of I have dimension at most n, the same holds for all irreducible representations π of A with $\ker(\pi)$ contained in $\overline{\operatorname{Prim}(I)}$. So by the correspondence described in (ii), irreducible representations of $\tau(B)$ are also at most n-dimensional. Hence, in this case, finitenesss of $\partial \operatorname{Prim}(I)$ is equivalent to finite-dimensionality of $\tau(B)$. ☐

For technical reasons we would prefer to work with unital extensions. However, it is not clear whether unitization preserves the regularity of R, i.e. whether the retract map associated to a unitized extension $0 \to \mathcal{C}_0(X, \mathbb{M}_N) \to A^+ \to B^+ \to 0$ is lower semicontinuous provided that the retract map associated to the original extension is. As the next lemma shows, this is true and holds in fact for more general extensions.

LEMMA 4.1.6. *Let a locally compact space X with clopen connected components and a commutative diagram*

$$\begin{tikzcd}
 & & 0 \arrow[d] & 0 \arrow[d] & \\
0 \arrow[r] & \mathcal{C}_0(X,\mathbb{M}_N) \arrow[r] \arrow[d, equal] & A \arrow[r] \arrow[d] & B \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & \mathcal{C}_0(X,\mathbb{M}_N) \arrow[r] & C \arrow[r] \arrow[d, "\pi"] & D \arrow[r] \arrow[d] & 0 \\
 & & F \arrow[r, equal] \arrow[d] & F \arrow[d] & \\
 & & 0 & 0 &
\end{tikzcd}$$

of short exact sequences of separable C^*-algebras be given. Let $R\colon \operatorname{Prim}(A) \to 2^{\operatorname{Prim}(B)}$ (resp. $S\colon \operatorname{Prim}(C) \to 2^{\operatorname{Prim}(D)}$) be the set-valued retract map associated to the upper (resp. the lower) horizontal sequence as in 4.1.1. If the quotient F is a finite-dimensional C^*-algebra, then the following holds:

1. *If R has pointwise finite image, then so does S.*
2. *If R is lower semicontinuous, then so is S.*

Proof. (1) This is immediate since $\operatorname{Prim}(F)$ is a finite set and one easily verifies $S(x) \subseteq R(x) \cup \operatorname{Prim}(F)$ for all $x \in X$.

(2) We may assume that F is simple and hence π is irreducible. Note that $S(J) = R(J)$ for all $J \in \operatorname{Prim}(B) \subset \operatorname{Prim}(D)$, while for $x \in X$ we have either $S(x) = R(x)$ or $S(x) = R(x) \cup \{[\pi]\}$. Given a closed subset $K \subseteq \operatorname{Prim}(D)$, we need to verify that $\{J \in \operatorname{Prim}(C)\colon S(J) \subseteq K\}$ is closed in $\operatorname{Prim}(C)$. If $[\pi] \in K$, then $\{J \in \operatorname{Prim}(C)\colon S(J) \subseteq K\} = \{J \in \operatorname{Prim}(A)\colon R(J) \subseteq K\} \cup \{[\pi]\}$ is closed in $\operatorname{Prim}(C)$ because $\{J \in \operatorname{Prim}(A)\colon R(J) \subseteq K\}$ is closed in $\operatorname{Prim}(A)$ by semicontinuity of R. Now if $[\pi] \notin K$, the only relevant case to check is a sequence $x_n \subset X$ converging to $\overline{x} \in \operatorname{Prim}(D)$ with $S(x_n) \subseteq K$ for all n. We then need to show that $S(\overline{x}) = \overline{x} \in K$ as well. Decompose $X = \bigcup_{i \in I} X_i$ into its clopen connected components and write $x_n \in X_{i_n}$ for suitable $i_n \in I$. We may assume that $i_n \neq i_m$ for $n \neq m$ since otherwise $\overline{x} \in \partial X_{i_n} = S(x_n)$ for some

n. Since R is lower semicontinuous, we know that the boundary of $\bigcup_n X_{i_n}$ in $\mathrm{Prim}(A)$ is contained in $K \cap \mathrm{Prim}(A)$ and hence $\partial\left(\bigcup_n X_{i_n}\right) \subset K \cup \{[\pi]\}$ in $\mathrm{Prim}(C)$.
Let p denote the projection of $\mathcal{C}_0(X, \mathbb{M}_N)$ onto $\mathcal{C}_0(\bigcup_n X_{i_n}, \mathbb{M}_N)$. This map canonically extends to $\overline{p}$ and $\overline{\overline{p}}$ making the diagram

$$\begin{tikzcd}
0 \arrow[r] & \mathcal{C}_0(X,\mathbb{M}_k) \arrow[r] \arrow[d, "p"] & C \arrow[r] \arrow[d, "\overline{p}"] & D \arrow[r] \arrow[d, "\overline{\overline{p}}"] & 0 \\
0 \arrow[r] & \bigoplus\limits_n \mathcal{C}_0(X_{i_n},\mathbb{M}_N) \arrow[r] \arrow[d, "\subseteq"] & \prod\limits_n \mathcal{C}(\beta X_{i_n},\mathbb{M}_N) \arrow[r] \arrow[d, equal] & \frac{\prod_n \mathcal{C}(\beta X_{i_n},\mathbb{M}_N)}{\bigoplus_n \mathcal{C}_0(X_{i_n},\mathbb{M}_N)} \arrow[r] \arrow[d, "q"] & 0 \\
0 \arrow[r] & \prod\limits_n \mathcal{C}_0(X_{i_n},\mathbb{M}_N) \arrow[r] & \prod\limits_n \mathcal{C}(\beta X_{i_n},\mathbb{M}_N) \arrow[r] & \prod\limits_n \mathcal{C}(\chi(X_{i_n}),\mathbb{M}_N) \arrow[r] & 0
\end{tikzcd}$$

commute. Using Lemma 4.1.5, we can indentify the boundary of $\bigcup_n X_{i_n}$ in $\mathrm{Prim}(C)$ with $\mathrm{Prim}\left(\overline{\overline{p}}(D)\right)$. We already know that $\overline{\overline{p}}$ factors through $D_K \oplus F$, where D_K denotes the quotient corresponding to the closed subset K of $\mathrm{Prim}(D)$, and denote the induced map by φ:

$$\begin{tikzcd}
D \arrow[rr, "\overline{\overline{p}}"] \arrow[dr, "\pi_K \oplus \pi"'] & & \frac{\prod_n \mathcal{C}(\beta X_{i_n},\mathbb{M}_N)}{\bigoplus_n \mathcal{C}_0(X_{i_n},\mathbb{M}_N)} \\
& D_K \oplus F \arrow[ur, dashed, "\varphi"'] &
\end{tikzcd}$$

We further know that the composition $q \circ \varphi_{|F} \colon F \to \prod_n \mathcal{C}(\chi(X_{i_n}), \mathbb{M}_N)$ vanishes because $[\pi] \notin \partial X_{i_n} = R(x_n) \subseteq K$ for all n. Hence the image of F under φ is contained in $\ker(q) = \frac{\prod_n \mathcal{C}_0(X_{i_n},\mathbb{M}_N)}{\bigoplus_n \mathcal{C}_0(X_{i_n},\mathbb{M}_N)}$. But since this C^*-algebra is projectionless and F is finite-dimensional, we find $\varphi_{|F} = 0$. Consequently, $\overline{\overline{p}}$ factors through D_K which means nothing but $\overline{x} \in \partial\left(\bigcup_n X_{i_n}\right) = \mathrm{Prim}\left(\overline{\overline{p}}(D)\right) \subseteq K$. ☐

Lemma 4.1.7. *Let X be a connected, compact space of dimension at most 1. For every finite-dimensional C^*-algebra $F \subseteq \mathcal{C}(X, \mathbb{M}_n)$ there exists a unitary $u \in \mathcal{C}(X, \mathbb{M}_n)$ such that uFu^* is contained in the constant $\mathbb{M}_n$-valued functions on X.*

Proof. Since $\dim(X) \leq 1$, equivalence of projections in $\mathcal{C}(X, \mathbb{M}_n)$ is completely determined by their rank ([Phi07, Proposition 4.2]). In particular, the C^*-algebra $\mathcal{C}(X, \mathbb{M}_n)$ has cancellation. Hence [RLL00, Lemma 7.3.2] shows that the inclusion $F \subset \mathcal{C}(X, \mathbb{M}_n)$ is unitarily equivalent to any constant embedding $\iota \colon F \to \mathbb{M}_n \subseteq \mathcal{C}(X, \mathbb{M}_n)$ with $\mathrm{rank}(\iota(p)) = \mathrm{rank}(p)$ for all minimal projections $p \in F$. ☐

Lemma 4.1.8. *Let X be a connected, locally compact, metrizable space of dimension at most 1. Then every unitary in $\mathcal{C}(\chi(X),\mathbb{M}_n)$ lifts to a unitary in $\mathcal{C}(\beta X,\mathbb{M}_n)$.*

Proof. By [Phi07, Proposition 4.2], we have $K_0(\mathcal{C}(\alpha X,\mathbb{M}_n)) \cong \mathbb{Z}$ via $[p] \mapsto \mathrm{rank}(p)$. Using the 6-term exact sequence in K-theory, this shows that the induced map $K_1(\mathcal{C}(\beta X,\mathbb{M}_n)) \to K_1(\mathcal{C}(\chi(X),\mathbb{M}_n))$ is surjective. Combining this with K_1-bijectivity of $\mathcal{C}(\beta X,\mathbb{M}_n)$, which is guaranteed by $\dim(\beta X) = \dim(X) \leq 1$ ([Nag70, Thm. 9.5]) and [Phi07, Theorem 4.7], the claim follows. $\square$

Proposition 4.1.9. *Let a short exact sequence of separable C^*-algebras*

$$0 \longrightarrow \mathcal{C}_0(X,\mathbb{M}_N) \longrightarrow A \longrightarrow B \longrightarrow 0 \qquad [\tau]$$

with Busby invariant τ be given. Assume that X is at most one-dimensional, has clopen connected components $(X_i)_{i\in I}$ and that every corona space $\chi(X_i)$ has only finitely many connected components. If the associated set-valued retract map R as in 4.1.1 has pointwise finite image, then there is a unitary $U \in \mathcal{C}(\beta X,\mathbb{M}_N)$ such that for each $i \in I$ the composition

$$B \xrightarrow{\tau} \mathcal{C}(\chi(X),\mathbb{M}_n) \xrightarrow{\mathrm{Ad}(\varrho(U))} \mathcal{C}(\chi(X),\mathbb{M}_N) \to \mathcal{C}(\chi(X_i),\mathbb{M}_n)$$

has image contained in the locally constant $\mathbb{M}_N$-valued functions on $\chi(X_i)$.

Proof. By Lemma 4.1.5, the image of each $\tau_i\colon B \xrightarrow{\tau} \mathcal{C}(\chi(X),\mathbb{M}_N) \to \mathcal{C}(\chi(X_i),\mathbb{M}_N)$ is finite-dimensional. Since by [Nag70, Thm. 9.5] furthermore $\dim\chi(X_i) \leq \dim\beta X_i = \dim X_i \leq \dim X \leq 1$, we can apply Lemma 4.1.7 to obtain unitaries $u_i \in \mathcal{C}(\chi(X_i),\mathbb{M}_N)$ such that $u_i\tau_i(B)u_i^*$ is contained in the locally constant functions on $\chi(X_i)$. These unitaries can be lifted to unitaries $U_i \in \mathcal{C}(\beta X_i,\mathbb{M}_N)$ by Lemma 4.1.8. Now $U = \oplus_i U_i \in \prod_i \mathcal{C}(\beta X_i,\mathbb{M}_N) = \mathcal{C}(\beta X,\mathbb{M}_N)$ has the desired property. $\square$

Lemma 4.1.10. *Let a short exact sequence of separable C^*-algebras*

$$0 \longrightarrow \mathcal{C}_0(X,\mathbb{M}_N) \longrightarrow A \longrightarrow B \longrightarrow 0 \qquad [\tau]$$

with Busby map τ be given. Assume that X is at most one-dimensional and that the connected components $(X_i)_{i\in I}$ of X are clopen. Further assume that the image of τ is constant on each $\chi(X_i) \subseteq \chi(X)$. Denote by $\iota\colon A \to \mathcal{M}(\mathcal{C}_0(X,\mathbb{M}_N)) = \mathcal{C}(\beta X,\mathbb{M}_N)$ the canonical map. If the set-valued retract map $R\colon \mathrm{Prim}(A) \to 2^{\mathrm{Prim}(B)}$ as defined in 4.1.1 is lower semicontinuous, the following statement holds:
For every finite set $\mathcal{G} \subset A$, every $\epsilon > 0$ and almost every $i \in I$ there exists a unitary $U_i \in \mathcal{C}(\alpha X_i,\mathbb{M}_N) \subset \mathcal{C}(\beta X_i,\mathbb{M}_N)$ such that

$$\left\|(U_i\iota(a)_{|\beta X_i}U_i^*)(x) - \iota(a)(y)\right\| < \epsilon$$

holds for all $a \in \mathcal{G}$, $x \in \beta X_i$ and $y \in \chi(X_i)$.

Proof. We may assume that A is unital by Lemma 4.1.6. Let a finite set $\mathcal{G} \subset A$ and $\epsilon > 0$ be given. For each $x \in \beta X$, we write $F_x = \mathrm{im}(\mathrm{ev}_x \circ \iota) \subseteq \mathbb{M}_N$ and

$$T_1(F_x) = \{f \in \mathcal{C}([0,1], F_x) \colon f(0) \in \mathbb{C} \cdot 1_{F_x}\},$$

$$S_1(F_x) = \{f \in \mathcal{C}_0([0,1), F_x) \colon f(0) \in \mathbb{C} \cdot 1_{F_x}\}.$$

Further let $h_\eta \in \mathcal{C}_0[0,1)$ denote the function $t \mapsto \max\{1-t-\eta, 0\}$. Using the Urysohn-type result [ELP98, Theorem 2.3.3], we find for each $x \in \beta X$ a commuting diagram

$$\begin{tikzcd}
0 \arrow[r] & J_x \arrow[r] & A \arrow[r, "\mathrm{ev}_x \circ \iota"] & F_x \arrow[r] & 0 \\
0 \arrow[r] & S(\mathbb{C}, F_x) \arrow[r] \arrow[u, "\varphi_x"] & T(\mathbb{C}, F_x) \arrow[r, "\mathrm{ev}_1"'] \arrow[u, "\overline{\varphi}_x"] & F_x \arrow[r] \arrow[u, equal] \arrow[l, dotted, bend right, "s_x"'] & 0
\end{tikzcd}$$

such that $\overline{\varphi}_x$ is unital and φ_x is proper. Let $s_x \colon F_x \to T(\mathbb{C}, F_x)$ be any map satisfying $s_x(b)(t) = b$ for $t \geq 1/2$, so that in particular $\mathrm{ev}_1 \circ s_x = \mathrm{id}_{F_x}$ holds. Now consider

$$V_{x,\delta} = \{y \in \beta X \colon \quad (\mathrm{ev}_y \circ \iota)(\overline{\varphi}_x(h_\delta)) = 0\}$$

which is, for $\delta > 0$, a closed neighborhood of x in βX. Note that by assumption $\chi(X_i) \cap V_{x,\delta} \neq \emptyset$ implies $\chi(X_i) \subseteq V_{x,\delta}$. We further claim the following: For almost every $i \in I$ the inclusion $\chi(X_i) \subseteq V_{x,\delta}$ implies $X_i \subset V_{x,2\delta}$. Assume otherwise, then we find pairwise different $i_n \in I$, points $x_n \in X_{i_n}$ such that $\chi(X_{i_n}) \subseteq V_{x,\delta}$ while $x_n \notin V_{x,2\delta}$ for all n. We may assume that $\mathrm{ev}_{x_n} \circ \iota$ converges pointwise to a representation π. Then

$$\|\pi(\overline{\varphi}_x(h_\delta))\| = \lim_n \|(\mathrm{ev}_{x_n} \circ \iota \circ \overline{\varphi}_x)(h_\delta)\| \geq \delta$$

since $x_n \neq V_{x,2\delta}$ implies that $\mathrm{ev}_{x_n} \circ \iota \circ \overline{\varphi}$ contains irreducible summands corresponding to evaluations at points t with $t < 1 - 2\delta$. On the other hand, since the retract map R is lower semicontinuous, we find each irreducible summand of π to be the limit of irreducible subrepresentations ϱ_n of $\mathrm{ev}_{y_n} \circ \iota$ where $y_n \in \chi(X_{i_n}) \subseteq V_{x,\delta}$. Hence

$$\|\pi(\overline{\varphi}_x(h_\delta))\| \leq \liminf_n \|\varrho_n(\overline{\varphi}_x(h_\delta))\| = 0$$

by 2.2.4, giving a contradiction and thereby proving our claim.
Since φ_x is proper, we have $J_x = \overline{\bigcup_{\eta > 0} \mathrm{her}(\varphi_x(h_\eta))}$. Hence there exists $1/2 > \delta(x) > 0$ such that

$$\inf\left\{\|(a - (\overline{\varphi}_x \circ s_x \circ \mathrm{ev}_x \circ \iota)(a)) - b\| \colon b \in \mathrm{her}(\varphi_x(h_{2\delta(x)}))\right\} < \frac{\epsilon}{2}$$

for all $a \in \mathcal{G}$. By compactness of $\chi(X)$, we find $x_1, ..., x_m$ such that

$$\chi(X) \subseteq \bigcup_{j=1}^m V_{x_j, \delta(x_j)}.$$

Then by the claim proved earlier, for almost every i with $\chi(X_i) \subseteq V_{x_j,\delta(x_j)}$ we have a factorization as indicated

$$\begin{tikzcd} A \arrow[r, "\iota"] \arrow[d, "\mathrm{ev}_{x_j} \circ \iota"'] \arrow[dr, "\pi_j"] & \prod_i \mathcal{C}(\alpha X_i, \mathbb{M}_N) \arrow[r] & \mathcal{C}(\alpha X_i, \mathbb{M}_N), \\ F_{x_j} \arrow[r, "\pi_j \circ \overline{\varphi}_{x_j} \circ s_{x_j}"'] & A/\langle \overline{\varphi}_{x_j}(h_{2\delta(x_j)}) \rangle \arrow[ur, dashed, "\overline{\iota}_i"'] & \end{tikzcd}$$

where $\langle \overline{\varphi}_{x_j}(h_{2\delta(x_j)}) \rangle$ denotes the ideal generated by $\overline{\varphi}_{x_j}(h_{2\delta(x_j)})$ and π_j the corresponding quotient map. By the choice of $\delta(x_j)$, the lower left triangle commutes up to $\epsilon/2$ on the finite set $\mathcal{G}$. Also note that the map $\pi_j \circ \overline{\varphi}_{x_j} \circ s_{x_j}$ is multiplicative.
Finally, by Lemma 4.1.7 there exists a unitary $U_i \in \mathcal{C}(\alpha X_i, \mathbb{M}_N)$ such that $\mathrm{Ad}(U_i) \circ (\overline{\iota}_i \circ \pi_j \circ \overline{\varphi}_{x_j} \circ s_{x_j})$ is a constant embedding. Of course, we may arrange $U(\infty) = 1$. We then verify

$$\begin{array}{cl} & \|(U_i \iota(a)_{|\beta X_i} U_i^*)(x) - \iota(a)(y)\| \\ \leq & \|(U_i(\overline{\iota}_i \circ \pi_j \circ \overline{\varphi}_{x_j} \circ s_{x_j})((\mathrm{ev}_{x_j} \circ \iota)(a))U_i^*)(x) - (\overline{\iota}_i \circ \pi_j)(a)(y)\| + \frac{\epsilon}{2} \\ \leq & \|(U_i(\overline{\iota}_i \circ \pi_j \circ \overline{\varphi}_{x_j} \circ s_{x_j})((\mathrm{ev}_{x_j} \circ \iota)(a))U_i^*)(y) - (\overline{\iota}_i \circ \pi_j)(a)(y)\| + \frac{\epsilon}{2} \\ = & \|(\overline{\iota}_i \circ (\pi_j \circ \overline{\varphi}_{x_j} \circ s_{x_j}) \circ (\mathrm{ev}_{x_j} \circ \iota))(a)(y) - (\overline{\iota}_i \circ \pi_j)(a)(y)\| + \frac{\epsilon}{2} \\ \leq & \epsilon. \end{array}$$

Applying this procedure to each of the finitely many points $x_1, ..., x_m$, the statement of the lemma follows. $\square$

Using Lemma 4.1.10 we can now construct a split for our sequence of interest - at least in the case of $\tau(B)$ being constant on each $\chi(X_i)$.

Corollary 4.1.11. *If* $0 \to \mathcal{C}_0(X, \mathbb{M}_N) \to A \to B \to 0$ *is a short exact sequence of separable C^*-algebras such that the assumptions of Lemma 4.1.10 hold, then this sequence splits.*

Proof. Let $\tau\colon B \to \mathcal{Q}(\mathcal{C}_0(X, \mathbb{M}_N)) = \mathcal{C}(\chi(X), \mathbb{M}_N)$ denote the Busby map of the sequence. We have the canonical commutative diagram

$$\begin{tikzcd} 0 \arrow[r] & \bigoplus_i \mathcal{C}_0(X_i, \mathbb{M}_N) \arrow[r] & \mathcal{C}(\beta X, \mathbb{M}_N) \arrow[r, "\varrho"] & \mathcal{C}(\chi(X), \mathbb{M}_N) \arrow[r] & 0 \\ 0 \arrow[r] & \mathcal{C}_0(X, \mathbb{M}_N) \arrow[r] \arrow[u, equal] & A \arrow[r, "\pi"] \arrow[u, "\iota"] & B \arrow[r] \arrow[u, "\tau"] & 0. \end{tikzcd}$$

Choose points $y_i \in \chi(X_i)$ for every $i \in I$. Using separability of A and Lemma 4.1.10, we find a unitary $U \in \prod_i \mathcal{C}(\alpha X_i, \mathbb{M}_N) \subset \prod_i \mathcal{C}(\beta X_i, \mathbb{M}_N) = \mathcal{C}(\beta X, \mathbb{M}_N)$ with

$$U\iota(a)U^* - \prod_i \iota(a)(y_i) \cdot 1_{\alpha X_i} \in \bigoplus_i \mathcal{C}_0(X_i, \mathbb{M}_N)$$

for all $a \in A$ (where $\iota(a)(y_i) \cdot 1_{\alpha X_i}$ denotes the function on αX_i with constant value $\iota(a)(y_i)$). By setting $s(\pi(a)) = U^* \left(\prod_i (\iota(a)(y_i) \cdot 1_{\alpha X_i}\right) U$ we find $s\colon B \to \mathcal{C}(\beta X, \mathbb{M}_N)$ with $(\varrho \circ s)(\pi(a)) = (\varrho \circ \iota)(a) = \tau(\pi(a))$ by the formula above. Identifying A with the pullback over ϱ and τ, we can regard s as a map from B to A with $\pi \circ s = \mathrm{id}_B$, i.e. we have constructed a split for the sequence. □

As the example $0 \to \mathcal{C}_0(0,1) \to \mathcal{C}[0,1] \to \mathbb{C}^2 \to 0$ shows, we cannot expect extensions by $\mathcal{C}_0(X, M_N)$ to split if the corona space of X (or of one of its components) is not connected. We will now deal with these components and show that one can still obtain a split $s\colon B \to A$ which, though not multiplicative in general, has still good multiplicative properties.

Lemma 4.1.12. *Let $0 \to \mathcal{C}_0(X, \mathbb{M}_N) \to A \to B \to 0$ be a short exact sequence with Busby map τ. Assume that the corona space $\chi(X)$ of X has only finitely many connected components and that the image of τ is contained in the locally constant functions on $\chi(X)$. Then there exists a compact set $K \subset X$ and a completely positive split $s\colon B \to \mathcal{C}(\beta X, \mathbb{M}_N)$ which is multiplicative outside of an open set $U \subset K$.*

Proof. Let $\chi(X) = \bigcup_{k=1}^K Y_k$ be the decomposition of the corona space into its connected components. By assumption τ decomposes as $\oplus_{k=1}^K \tau_k$ with $\mathrm{im}(\tau_k) \subset \mathbb{M}_N \cdot 1_{Y_k} \subseteq \mathcal{C}(\chi(X), \mathbb{M}_N)$. Lift the indicator functions $1_{Y_1}, \cdots, 1_{Y_K}$ to pairwise orthogonal contractions $h_1, \cdots, h_K$ in $\mathcal{C}(\beta X, \mathbb{C} \cdot 1_{M_N})$ and let $f\colon [0,1] \to [0,1]$ be the continuous function which equals 1 on $\left[\frac{1}{2}, 1\right]$, satisfies $f(0) = 0$ and is linear in between. We define a completely positive map $s\colon B \to \mathcal{C}(\beta X, \mathbb{M}_N)$ by $s(b)(x) = \sum_{k=1}^K \tau_k(b) \cdot f(h_k)(x)$ and check that in the diagram

$$\begin{tikzcd}
A \arrow[r, "\iota"] \arrow[d] & \mathcal{C}(\beta X, \mathbb{M}_N) \arrow[d] \\
B \arrow[r, "\tau"] \arrow[ur, dashed, "s"] & \mathcal{C}(\chi(X), \mathbb{M}_N)
\end{tikzcd}$$

the right triangle commutes. Set $K = \bigcap_{k=1}^K h_k^{-1}([0, \frac{1}{2}]) \subset X$, then s is multiplicative outside of the open set $U = \bigcap_{k=1}^K h_k^{-1}([0, \frac{1}{2})) \subset K \subset X$. □

Proposition 4.1.13. *Let $0 \to \mathcal{C}_0(X, \mathbb{M}_N) \to A \to B \to 0$ be a short exact sequence of separable C^*-algebras with Busby map τ. Assume that X is at most one-dimensional and has clopen connected components $(X_i)_{i \in I}$. Further assume that each corona space $\chi(X_i)$ has only finitely many connected components and that $\chi(X_i)$ is connected for almost all $i \in I$. If for each $i \in I$ the image of $\tau_i\colon B \xrightarrow{\tau} \mathcal{C}(\chi(X), \mathbb{M}_N) \to \mathcal{C}(\chi(X_i), \mathbb{M}_N)$ is locally constant on $\chi(X_i)$ and the set-valued retract map $R\colon \mathrm{Prim}(A) \to 2^{\mathrm{Prim}(B)}$ as in 4.1.1 is lower semicontinuous, the following holds: There exists a compact set $K \subset X$ and a completely positive split $s\colon B \to \mathcal{C}(\beta X, \mathbb{M}_N)$ which is multiplicative outside of an open set $U \subset K$.*

Proof. Let $I_0 \subseteq I$ be a finite set such that $\chi(X_i)$ is connected for every $i \in I_1 := I \backslash I_0$. We may then study the extensions ($* = 0$ or 1)

$$\begin{tikzcd}
0 \arrow[r] & \bigoplus\limits_{i\in I} \mathcal{C}_0(X_i,\mathbb{M}_N) \arrow[r] \arrow[d, "pr_{I_*}"] & A \arrow[r] \arrow[d, equal] & B \arrow[r] \arrow[d, "\varphi_*"] & 0 \\
0 \arrow[r] & \bigoplus\limits_{i\in I_*} \mathcal{C}_0(X_i,\mathbb{M}_N) \arrow[r] \arrow[d, equal] & A \arrow[r] \arrow[d, "\iota_*"] & A/\bigoplus\limits_{i\in I_*} \mathcal{C}_0(X_i,\mathbb{M}_N) \arrow[r] \arrow[d, "\tau_*"] \arrow[dl, dashed, "s_*"'] & 0 \\
0 \arrow[r] & \bigoplus\limits_{i\in I_*} \mathcal{C}_0(X_i,\mathbb{M}_N) \arrow[r] & \prod\limits_{i\in I_*} \mathcal{C}(\beta X_i,\mathbb{M}_N) \arrow[r] & \frac{\prod_{i\in I_*} \mathcal{C}(\beta X_i,\mathbb{M}_N)}{\bigoplus_{i\in I_*} \mathcal{C}_0(X_i,\mathbb{M}_N)} \arrow[r] & 0
\end{tikzcd}$$

with Busby maps τ_*. Denote the map $B \to A/\bigoplus_{i\in I_*} \mathcal{C}_0(X_i,\mathbb{M}_N)$ induced by the projection pr_{I_*} by φ_*. It is now easy to check that for $* = 1$ the short exact sequence in the middle row satisfies the assumptions of Lemma 4.1.10 and hence admits a splitting s_1 by Corollary 4.1.11. For $* = 0$, we apply Lemma 4.1.12 to obtain a compact set $K \subset \bigsqcup_{i\in I_0} X_i$ and a completely positive split s_0 which is multiplicative outside of an open set $U \subset K \subset \bigsqcup_{i\in I_0} X_i$. Setting $s = s_0 \circ \varphi_0 \oplus s_1 \circ \varphi_1$, we now get a split for the original sequence. In particular, $\varrho \circ s = \tau$ holds due to the commutative diagram

$$\begin{tikzcd}
0 \arrow[r] & \mathcal{C}_0(X,\mathbb{M}_N) \arrow[r] \arrow[d, equal] & A \arrow[r] \arrow[d, "\iota_0\oplus\iota_1"] & B \arrow[r] \arrow[d, "\tau_0\oplus\tau_1"'] \arrow[dl, dashed, "s"'] \arrow[dd, bend left, "\tau"] & 0 \\
0 \arrow[r] & \bigoplus\limits_{*=0,1} \bigoplus\limits_{i\in I_*} \mathcal{C}_0(X_i,\mathbb{M}_N) \arrow[r] \arrow[d, equal] & \bigoplus\limits_{*=0,1} \prod\limits_{i\in I_*} \mathcal{C}(\beta X_i,\mathbb{M}_N) \arrow[r] \arrow[d, equal] & \bigoplus\limits_{*=0,1} \frac{\prod_{i\in I_*} \mathcal{C}(\beta X_i,\mathbb{M}_N)}{\bigoplus_{i\in I_*} \mathcal{C}_0(X_i,\mathbb{M}_N)} \arrow[r] \arrow[d, equal] & 0 \\
0 \arrow[r] & \bigoplus\limits_{i\in I} \mathcal{C}_0(X_i,\mathbb{M}_N) \arrow[r] & \prod\limits_{i\in I} \mathcal{C}(\beta X_i,\mathbb{M}_N) \arrow[r, "\varrho"] & \frac{\prod_{i\in I} \mathcal{C}(\beta X_i,\mathbb{M}_N)}{\bigoplus_{i\in I} \mathcal{C}_0(X_i,\mathbb{M}_N)} \arrow[r] & 0.
\end{tikzcd}$$

□

Summarizing the results of this section, we obtain the following.

Theorem 4.1.14. *Let a short exact sequence of separable C^*-algebras*

$$0 \longrightarrow \mathcal{C}_0(X,\mathbb{M}_N) \longrightarrow A \longrightarrow B \longrightarrow 0 \qquad [\tau]$$

with Busby map τ be given. Assume that X satisfies the conditions

1. $\dim X \leq 1$,
2. *the connected components $(X_i)_{i\in I}$ of X are clopen,*
3. *each $\chi(X_i)$ has finitely many connected components,*

4. almost all $\chi(X_i)$ are connected,

then the following holds: If the associated set-valued retract map $R\colon \mathrm{Prim}(A) \to 2^{\mathrm{Prim}(B)}$ given as in 4.1.1 by

$$R(z) = \begin{cases} z \text{ if } z \in \mathrm{Prim}(B) \\ \partial X_i = \overline{X_i} \backslash X_i \text{ if } z \in X_i \subseteq X \end{cases}$$

is lower semicontinuous and has pointwise finite image, then there exists a compact set $K \subset X$ and a completely positive split $s\colon B \to A$ for the sequence such that the composition

$$B \xrightarrow{s} A \longrightarrow \mathcal{M}(\mathcal{C}_0(X, \mathbb{M}_N)) = \mathcal{C}_b(X, \mathbb{M}_N)$$

is multiplicative outside of an open set $U \subset K$.

Proof. Note that we can replace the given extension by any strongly unitarily equivalent one (in sense of [Bla06, II.8.4.12]) without changing the retract map R. Hence, by Proposition 4.1.9, we may assume that the image of τ is locally constant on each $\chi(X_i)$. Now Proposition 4.1.13 provides a split s with the desired properties. □

4.1.3 Retract maps for semiprojective extensions

We now verify the regularity properties for the set-valued retract map $R\colon \mathrm{Prim}(A) \to 2^{\mathrm{Prim}(B)}$ associated to an extension $0 \to \mathcal{C}_0(X, \mathbb{M}_N) \to A \to B \to 0$ in the case that both the ideal $\mathcal{C}_0(X, \mathbb{M}_N)$ and the extension A are semiprojective C^*-algebras.

First we need the following definition which is an adaption of 2.1.2 and 2.1.3 to the setting of pointed spaces.

Definition 4.1.15. *Let (X, x_0) be a pointed one-dimensional Peano continuum and $r\colon X \to \mathrm{core}(X)$ the first point map onto the core of X as in 2.1.3 (where we choose $\mathrm{core}(X)$ to be any point $x \neq x_0$ if X is contractible). Denote the unique arc from x_0 to $r(x_0)$ by $[x_0, r(x_0)]$, then we say that*

$$\mathrm{core}(X, x_0) := \mathrm{core}(X) \cup [x_0, r(x_0)]$$

is the core of (X, x_0). It is the smallest subcontinuum of X which contains both $\mathrm{core}(X)$ and the point x_0.

Now let X be a non-compact space with the property that its one-point compactification $\alpha X = X \cup \{\infty\}$ is a one-dimensional ANR-space. We are interested in the structure of the space X at around infinitity (i.e. outside of large compact sets) which is reflected in its corona space $\chi(X) = \beta X \backslash X$. At least some information about $\chi(X)$ can be obtained by studying neighborhoods of the point ∞ in αX. The following lemma describes some special neighborhoods which relate nicely to the finite graph $\mathrm{core}(\alpha X, \infty)$.

LEMMA 4.1.16. *Let X be a connected, non-compact space such that its one-point compactification $\alpha X = X \cup \{\infty\}$ is a one-dimensional ANR-space. Fix a geodesic metric d on αX (which exists by [Bin49], [Moi49]), then for any compact set $C \subset \alpha X \backslash \{\infty\}$ there exists a closed neighborhood V of ∞ with the following properties:*

(i) $\{x \in X\colon d(x,\infty) \leq \epsilon\} \subseteq V \subseteq X\backslash C$ *for some* $\epsilon > 0$.

(ii) $V \cap \operatorname{core}(\alpha X, \infty)$ *is homeomorphic to the space of* K *many intervals* $[0,1]$ *glued together at the* 0*-endpoints with* $K = \operatorname{order}(\infty, \operatorname{core}(\alpha X, \infty))$*. The gluing point corresponds to* ∞ *under this identification.*

Let $D^{(k)} \subseteq V$ *denote the* k*-th copy of* $[0,1]$ *under the identification described above and let* r *be the first point map onto* $\operatorname{core}(\alpha X, \infty)$*. We can further arrange:*

(iii) $V = \bigcup_{k=1}^{K} r^{-1}\left(D^{(k)}\right)$ *and* $r^{-1}\left(D^{(k\ )}\right) \cap r^{-1}\left(D^{(k')}\right) = \{\infty\}$ *for* $k \neq k'$.

(iv) The connected components of $V\backslash\{\infty\}$ *are given by* $V^{(k)} := r^{-1}\left(D^{(k)}\backslash\{\infty\}\right)$.

(v) Every path in V *from* $x \in V^{(k)}$ *to* $x' \in V^{(k')}$ *with* $k \neq k'$ *contains* ∞.

Proof. We first note that $r^{-1}(\{\infty\}) \cap X$ is open. Assume there is $x \in X$ with $r(x) = \infty$ and $d(x,\infty) = r > 0$. Then given any $y \in X$ with $d(x,y) < r$ we choose an isometric arc $\alpha\colon [0, d(x,y)] \to \alpha X$ from x to y. Now the arc from y to ∞ given by first following α in reverse direction and then going along the unique arc from x to ∞ must run through $r(y)$ by 2.1.3. Since every point on the second arc gets mapped to ∞ by r, we find either $r(y) = \infty$ or there is $0 < t < d(x,y)$ such that $\alpha(t) = r(y) \in \operatorname{core}(\alpha X, \infty)$. In the second case, the arc $\alpha_{|[0,t]}$ must run through $r(x) = \infty$ which, using the fact that α was isometric, gives the contradiction $d(x,\infty) < t < d(x,y) < r$. Since $r^{-1}(\{\infty\})$ is also closed, connectedness of X implies in fact $r^{-1}(\{\infty\}) = \{\infty\}$.
By definition of K (see section 2.1), the closed set $\{x \in \operatorname{core}(\alpha X, \infty)\colon d(x,\infty) \leq \epsilon\}$ satisfies the description in (ii) for all sufficiently small $\epsilon > 0$. We set

$$V = \{x \in \alpha X\colon d(r(x), \infty) \leq \epsilon\},$$

then $V \cap \operatorname{core}(\alpha X, \infty) = r(V) = \{x \in \operatorname{core}(\alpha X, \infty)\colon d(x,\infty) \leq \epsilon\}$ so that condition (ii) is satisfied. For (i), we observe that $d(x,\infty) \leq \epsilon$ implies $d(r(x),\infty) \leq \epsilon$ since d is geodesic and every arc from x to ∞ runs through $r(x)$. Since $\infty \notin r(C)$, we have $\min\{d(r(x),\infty)\colon x \in C\} > 0$ and therefore $V \cap C = \emptyset$ for ϵ sufficiently small. Condition (iii) follows immediately from the definition of V. The sets $V^{(k)}$ are connected and open by construction, so that (iv) holds. (v) follows from (iv). ◻

We now collect some information about the corona space $\chi(X)$ in the case of connected X. These observations are mostly based on the work of Grove and Pedersen in [GP84] and the graph-like structure of one-dimensional ANR-spaces.

Lemma 4.1.17. *Let X be a connected, non-compact space such that its one-point compactification αX is a one-dimensional ANR-space. Then the corona space $\chi(X)$ has covering dimension at most 1 and its number of connected components is given by $K = \mathrm{order}(\infty, \mathrm{core}(\alpha X, \infty)) < \infty$. In particular, if αX is a one-dimensional AR-space, then $\chi(X)$ is connected.*

Proof. Apply Lemma 4.1.16 to $(\alpha X, \infty)$. It is straightforward to check that the map

$$\mathcal{C}(\chi(X)) = \mathcal{C}_b(X)/\mathcal{C}_0(X) \to \bigoplus_{k=1}^{K} \mathcal{C}_b(V^{(k)})/\mathcal{C}_0(V^{(k)}) = \bigoplus_{k=1}^{K} \mathcal{C}(\chi(V^{(k)}))$$

is an isomorphism. Therefore we find $\chi(X) = \bigsqcup_{k=1}^{K} \chi(V^{(k)})$ and it suffices to check that each $\chi(V^{(k)})$ is connected. By Proposition 3.5 of [GP84], it is now enough to show that each $V^{(k)}$ is connected at infinity. So let a compact set $C_1 \subset V^{(k)}$ be given and denote by $r\colon V^{(k)} \cup \{\infty\} \to D^{(k)}$ the first point map. Using the identification $[0,1] \cong D^{(k)}$ where the point 0 corresponds to the point ∞, we find $t > 0$ such that $r(C_1) \subset [t,1]$. But $C_2 := r^{-1}([t,1])$ is easily seen to be compact while $V^{(k)} \backslash C_2 = r^{-1}((0,t))$ is path-connected by definition of r. For the dimension statement we note that $\dim(\chi(X)) \leq \dim(\beta X) = \dim(X) \leq 1$ by [Nag70, Theorem 9.5]. □

Remark 4.1.18. *The assumption that X is connected in 4.1.17 is necessary. If we drop it, the corona space $\chi(X)$ may no longer have finitely many connected components, but the following weaker statement holds: If αX is a one-dimensional ANR-space, so will be αX_i for any connected component X_i of X. However, it follows from 2.1.7 that all but finitely many components lead to contractible spaces αX_i, i.e. to one-dimensional AR-spaces. Since in this case $\mathrm{core}(\alpha X_i, \infty)$ is just an arc $[x, \infty]$ for some $x \in X_i$, we see from Lemma 4.1.17 that $\chi(X_i)$ is connected for almost every component X_i of X.*

We will now see that, in the situation described in the beginning of this section, the set-valued retract map R has pointwise finite image, i.e. $|R(z)| < \infty$ for all $z \in \mathrm{Prim}(A)$. The cardinality of these sets is in fact uniformly bounded and we give an upper bound which only depends on N and the structure of the finite graph $\mathrm{core}(\alpha X, \infty)$.

Proposition 4.1.19. *Let A be a semiprojective C^*-algebra containing an ideal of the form $\mathcal{C}_0(X, \mathbb{M}_N)$. If $\alpha X = X \cup \{\infty\}$ is a one-dimensional ANR-space, then every connected component C of X has finite boundary $\partial C = \overline{C} \backslash C$ in $\mathrm{Prim}(A)$. More precisely, we find*

$$|\partial C| \leq N \cdot \mathrm{order}(\infty, (\alpha C, \infty)) < \infty.$$

Proof. Since X is locally connected, the connected components of X are clopen and αC is again a one-dimensional ANR-space for every component C of X. Hence we may assume that $C = X$. Fix a geodesic metric d on $\alpha X = X \cup \{\infty\}$ ([Bin49], [Moi49]) and let V be a neighborhood of ∞ as constructed in Lemma 4.1.16, satisfying $\{x \in \alpha X \colon d(x,\infty) \leq \epsilon\} \subseteq V$ for some $\epsilon > 0$. We further choose sequences $(x_n^{(k)})_n \subseteq D^{(k)} \backslash \{\infty\}$ converging to ∞ and write $x_\infty^{(1)} = \cdots = x_\infty^{(K)} = \infty$. By compactness of the unit ball in $\mathbb{M}_N$ and separability of A, we may assume that the representation

$$\pi^{(k)} \colon A \to \mathbb{M}_N, \quad a \mapsto \lim_{n\to\infty} a(x_n^{(k)})$$

exists for all $1 \leq k \leq K$. Here, $a(x)$ denotes the image of $a \in A$ under the extension of the point evaluation $\mathrm{ev}_x \colon \mathcal{C}_0(X, \mathbb{M}_N) \to \mathbb{M}_N$ to A. For a sequence $(x_n)_n$ in $X \subseteq \mathrm{Prim}(A)$ we write $\mathrm{Lim}(x_n) = \{z \in \mathrm{Prim}(A) \colon x_n \to z\}$. Our goal is then to show that there exists a finite set $S \subset \mathrm{Prim}(A)$ such that $\mathrm{Lim}(x_n) \subset S$ for every sequence $(x_n)_n \subset X$ with $x_n \to \infty$ in αX. We will show that each $S^{(k)} := \mathrm{Lim}(x_n^{(k)})$ consists of at most N elements and that $S := \bigcup_{k=1}^K S^{(k)}$ has the desired property described above. First observe that

$$S^{(k)} = \left\{ \left[\pi_1^{(k)}\right], \ldots, \left[\pi_{r(k)}^{(k)}\right] \right\}$$

holds, where $\pi^{(k)} \simeq \pi_1^{(k)} \oplus \cdots \oplus \pi_{r(k)}^{(k)}$ is the decomposition of $\pi^{(k)}$ into irreducible summands. The $\supseteq$-inclusion is immediate, for the other direction assume that $x_n^{(k)} \to \ker(\varrho)$ for some irreducible representation ϱ with $\varrho \not\simeq \pi_i^{(k)}$ for all i. Since all $x_n^{(k)}$ correspond to N-dimensional representations, we also have $\dim(\varrho) \leq N$. Therefore all $\pi_i^{(k)}$ and ϱ drop to irreducible representations of the maximal N-subhomogeneous quotient $A_{\leq N}$ of A (cf. section 2.2). Because $\mathrm{Prim}(A_{\leq N})$ is a T_1-space, the finite set $\{[\pi_1^{(k)}], \ldots, [\pi_{r(k)}^{(k)}]\}$ is closed and $[\varrho]$ can be separated from it. In terms of 2.2.4, this means that there exists $a \in A$ such that $\|\varrho(a)\| > 1$ while $\|\pi_i^{(k)}(a)\| \leq 1$ for all i. On the other hand, we find

$$\|\varrho(a)\| \leq \liminf_{n\to\infty} \left\| a(x_n^{(k)}) \right\| = \left\| \pi^{(k)}(a) \right\| = \max_{i=1\ldots r(k)} \left\| \pi_i^{(k)}(a) \right\| \leq 1,$$

using 2.2.4 again. Hence $[\varrho] = [\pi_i^{(k)}]$ for some i and in particular $\left|S^{(k)}\right| = r(k) \leq N$ for every k.

It now suffices to show that $\mathrm{Lim}(x_n) \subseteq S^{(k)}$ for sequences $(x_n) \subset X$ with $x_n \to \infty$ such that $r(x_n) \in D^{(k)}$ for some fixed k and all n. Let such a sequence $(x_n)_n$ for some fixed k be given and pick $z \in \mathrm{Lim}(x_n)$. In order to show that $z \in S^{(k)}$, we consider the compact spaces

$$Y_n := \left\{ (t,t) | 0 \leq t \leq \frac{1}{n} \right\} \cup \bigcup_{m \geq n} \left(\left\{ \frac{1}{m} \right\} \times \left[0, \frac{1}{m} \right] \right) \subset \mathbb{R}^2.$$

Note that $Y_{n+1} \subset Y_n$ and $\bigcap_n Y_n = (0,0)$. We will now 'glue' $\mathcal{C}(Y_n, \mathbb{M}_N)$ to A in the following way: As before, we may assume that $x_n \to z$ in $\mathrm{Prim}(A)$ and that $\pi_\infty(a) = \lim_n a(x_n)$ exists for every $a \in A$. In particular, we find $z = [\pi_{i,\infty}]$ for some i where $\pi_\infty \simeq \pi_{1,\infty} \oplus \cdots \oplus \pi_{r_\infty,\infty}$ is the decomposition of π_∞ into irreducible summands. Let c denote the C^*-algebra of convergent $\mathbb{M}_N$-valued sequences, we can then form the pullback $A_n := A \oplus_c \mathcal{C}(Y_n, \mathbb{M}_N)$ over the two *-homomorphisms

$$\begin{gathered} A \longrightarrow c \\ a \mapsto (a(x_n), a(x_{n+1}), a(x_{n+2}), \dots) \end{gathered}$$

and

$$\begin{gathered} \mathcal{C}(Y_n, \mathbb{M}_N) \longrightarrow c. \\ f \mapsto f((\tfrac{1}{n}, 0), f(\tfrac{1}{n+1}, 0), f(\tfrac{1}{n+2}, 0), \dots) \end{gathered}$$

These pullbacks form an inductive system in the obvious way. Further note that the connecting maps $A_n \to A_{n+1}$ are all surjective. The limit $\varinjlim A_n$ can be identified with A via the isomorphism induced by the projections $A_n = A \oplus_c \mathcal{C}(Y_n, \mathbb{M}_N) \to A$ onto the left summand. Using semiprojectivity of A, we can find a partial lift to some finite stage A_n of this inductive system:

$$\begin{tikzcd} & & A_n = A \oplus_c \mathcal{C}(Y_n, \mathbb{M}_N) \arrow[r] \arrow[d] & \mathcal{C}(Y_n, \mathbb{M}_N) \\ \mathcal{C}_0(X, \mathbb{M}_N) \arrow[r, "\subseteq"] & A \arrow[ur, dashed] \arrow[r, "\cong"] & \varinjlim A_n & \end{tikzcd}$$

Let $\varphi\colon A \to \mathcal{C}(Y_n, \mathbb{M}_N)$ be the composition of this lift with the projection $A_n \to \mathcal{C}(Y_n, \mathbb{M}_N)$ to the right summand. The restriction of φ to the ideal $\mathcal{C}_0(X, \mathbb{M}_N)$ then induces a continuous map $\varphi^*\colon Y_n \to \alpha X$ with $\varphi^*\left(\frac{1}{m}, 0\right) = x_m$ for all $m \geq n$ and $\varphi^*(0,0) = \infty$. Denote by h the strictly positive element of $\mathcal{C}_0(X, \mathbb{M}_N)$ given by $h(x) = d(x, \infty) \cdot 1_{\mathbb{M}_N}$. After increasing n, we may assume that $\|\varphi(h)\| < \epsilon$ holds. For $m \geq n$, we consider the paths $\alpha_m\colon \left[0, \frac{2}{m}\right] \to Y_n$ given by

$$\alpha_m(t) = \begin{cases} \left(\frac{1}{m}, t\right) & \text{if } 0 \leq t \leq \frac{1}{m} \\ \left(\frac{2}{m} - t, \frac{2}{m} - t\right) & \text{if } \frac{1}{m} \leq t \leq \frac{2}{m}. \end{cases}$$

Set $t_{\infty,m} = \min\{t\colon \varphi(h)(\alpha_m(t)) = 0\}$, then $0 < t_{\infty,m} \leq \frac{2}{m}$ because of $\|\varphi(h)(\alpha_m(0))\| = \|\varphi(h)(\frac{1}{m}, 0)\| = \|h(x_m)\| = d(x_m, \infty) > 0$ and $\varphi(h)(\alpha_m(\frac{2}{m})) = \varphi(h)(0,0) = h(\infty) = 0$. By setting $\beta_m(t) = \varphi^*(\alpha_m(t))$ we obtain paths $\beta_m\colon [0, t_{\infty,m}] \to \alpha X$ which have the properties

(1) $\beta_m(0) = x_m$,
(2) $\beta_m(t) = \infty$ if and only if $t = t_{\infty,m}$,
(3) $\mathrm{im}(\beta_m) \subseteq V^{(k)}$ for all m,
(4) $x_l^{(k)} \in \mathrm{im}(\beta_m)$ for fixed m and all sufficiently large l.

The first property is clear while the second one follows directly from the definition of $t_{\infty,m}$. In order to verify properties (3) and (4) we have to involve the structure of the neighborhood V and by that the special structure of αX as a one-dimensional ANR-space. From $\|\varphi(h)\| < \epsilon$ we obtain $\mathrm{im}(\beta_m) \subseteq \mathrm{im}(\alpha_m) \subseteq \{x \in \alpha X \colon d(x,\infty) \leq \epsilon\} \subseteq V$, it then follows from (1), (2) and property (v) in Lemma 4.1.16 that $\mathrm{im}(\beta_m) \subseteq V^{(k)}$. For (4), observe that $\mathrm{im}(\beta_m)$ contains $r(\mathrm{im}(\beta_m))$ by part (ii) of Lemma 2.1.4, where r is the first-point map $\alpha X \to \mathrm{core}(\alpha X,\infty)$. Under the identification $D^{(k)} \cong [0,1]$, the connected set $r(\mathrm{im}(\beta_m))$ corresponds to a proper interval containing the 0-endpoint and hence it contains $x_l^{(k)}$ for almost every l.
Now set $\pi_m = \mathrm{ev}_{\beta(t_{\infty,m})} \circ \varphi \colon A \to \mathbb{M}_N$ and let $\pi_m \simeq \pi_{1,m} \oplus \cdots \oplus \pi_{r_m,m}$ be the decomposition into irreducible summands. We claim that the identity

$$S^{(k)} = \{[\pi_{1,m}], \cdots, [\pi_{r_m,m}]\}$$

holds for all m. Involving property (4) for the path β_m, we find

$$\|\pi_m(a)\| = \lim_{t \nearrow t_{\infty,m}} \left\|(\mathrm{ev}_{\beta(t)} \circ \varphi)(a)\right\| = \lim_{l\to\infty} \left\|a\left(x_l^{(k)}\right)\right\| = \left\|\pi^{(k)}(a)\right\|$$

for every fixed m and all $a \in A$. Now the same separation argument as in the beginning of the proof shows that the finite-dimensional representations $\pi^{(k)}$ and π_m share the same irreducible summands for every m. Since $\beta_m(t_{\infty,m}) \to (0,0)$ in Y_n, we find $\pi_m = \mathrm{ev}_{\beta(t_{\infty,m})} \circ \varphi \to \mathrm{ev}_{(0,0)} \circ \varphi = \pi_\infty$ pointwise. Hence by the above identity, π_∞ and $\pi^{(k)}$ also share the same irreducible summands. In particular, we find $z \in S^{(k)}$ which finishes the proof. □

Next, we show that in our situation the set-valued retract map R is also lower semicontinuous in the sense of 4.1.2.

PROPOSITION 4.1.20. *Let $0 \to \mathcal{C}_0(X,\mathbb{M}_N) \to A \to B \to 0$ be a short exact sequence of separable C^*-algebras. If αX is a one-dimensional ANR-space and A is semiprojective, then the associated retract map R as in 4.1.1 is lower semicontinuous.*

Proof. Let $X = \bigsqcup_{i\in I} X_i$ denote the decomposition of X into its connected components. By separability of A it suffices to verify condition (iii) of Lemma 4.1.2 for a given sequence $x_n \to z$ in $\mathrm{Prim}(A)$. The case $z \in X$ is trivial since X is locally connected and therefore has open connected components. The critical case is when $x_n \in X$ for all n but $z \in \mathrm{Prim}(B)$. In this case, we write $x_n \in X_{i_n}$ and we may assume that $\pi_\infty(a) = \lim_n a(x_n)$ is well defined for all $a \in A$. In particular, z corresponds to the kernel of an irreducible summand $\pi_{j,\infty}$ of $\pi_\infty \simeq \pi_{1,\infty} \oplus \cdots \oplus \pi_{r,\infty}$, as we have already seen in the beginning of the proof of Proposition 4.1.19. Using exactly the same construction of 'gluing the space Y to A along the sequence (x_n)' as in the proof of 4.1.19, one now shows that

$$\{[\pi_{1,\infty}], \cdots, [\pi_{r,\infty}]\} \subseteq \overline{\bigcup_n \partial X_{i_n}}.$$

Hence we find $y_n \in \partial X_{i_n} = R(x_n)$ with $y_n \to [\pi_{j,\infty}] = z$ showing that the retract map R is in fact lower semicontinuous. $\square$

4.2 Existence of limit structures

Consider an extension of separable C^*-algebras

$$0 \to \mathcal{C}_0(X, \mathbb{M}_N) \to A \to B \to 0$$

where the one-point compactification of X is assumed to be a one-dimensional ANR-space. We know from Theorem 2.1.5 that in this case αX comes as a inverse limit of finite graphs over a surprisingly simple system of connecting maps. Here we show that under the right assumptions on the set-valued retract map $R\colon \operatorname{Prim}(A) \to 2^{\operatorname{Prim}(B)}$ associated to the sequence above, this limit structure for αX is compatible with the extension of B by $\mathcal{C}_0(X, \mathbb{M}_N)$ in the following sense: We prove the existence of a direct limit structure for A which describes it as the C^*-algebra B with a sequence of non-commutative finite graphs (1-NCCW's) attached. The connecting maps of this direct system are obtained from the limit structure for αX and hence can be described in full detail.

Lemma 4.2.1. *Let a short exact sequence of separable C^*-algebras $0 \to \mathcal{C}_0(X, \mathbb{M}_N) \to A \to B \to 0$ with Busby map τ be given. Assume that αX is a one-dimensional ANR-space and that the associated set-valued retract map $R\colon \operatorname{Prim}(A) \to 2^{\operatorname{Prim}(B)}$ as in 4.1.1 is lower semicontinuous and has pointwise finite image. Then A is isomorphic to the direct limit $B_\infty = \varinjlim \left(B_i, s_i^{i+1}\right)$ of an inductive system*

$$\begin{tikzcd} B_0 \arrow[r, "s_0^1"'] & B_1 \arrow[r, "s_1^2"'] \arrow[l, bend right, "r_1^0"'] & B_2 \arrow[r] \arrow[l, bend right, "r_2^1"'] & \cdots \arrow[r, "s_{i-1}^i"'] \arrow[l, bend right] & B_i \arrow[r] \arrow[l, bend right, "r_i^{i-1}"'] \arrow[rr, bend right, "s_i^\infty"'] & \cdots \arrow[r] & B_\infty \arrow[ll, bend right, "r_\infty^i"'] \end{tikzcd}$$

where

- *B_0 is given as a pullback $B \oplus_F D$ with D a 1-NCCW and $\dim(F) < \infty$. Furthermore, if αX is contractible, we may even arrange $B_0 \cong B$.*

and

- *for every $i \in \mathbb{N}$ there is a representation $\pi_i\colon B_i \to \mathbb{M}_N$ such that B_{i+1} is defined as the pullback*
$$\begin{tikzcd} B_{i+1} \arrow[r, dotted] \arrow[d, dotted, "r_{i+1}^i"] & \mathcal{C}([0,1], \mathbb{M}_N) \arrow[d, "ev_0"] \\ B_i \arrow[r, "\pi_i"] \arrow[u, bend left, "s_i^{i+1}"] & \mathbb{M}_N. \end{tikzcd}$$
The map $s_i^{i+1}\colon B_i \to B_{i+1}$ is given by $a \mapsto (a, \pi_i(a) \otimes 1_{[0,1]})$ and hence satisfies $r_{i+1}^i \circ s_i^{i+1} = \mathrm{id}_{B_i}$.

Proof. Let $X = \bigsqcup_{j \in J} C_j$ be the decomposition of X into its clopen connected components. Denote by $J_1 \subseteq J$ the subset of those indices for which the corona space $\chi(C_j)$ is connected and note that $J_0 := J \backslash J_1$ is finite by Remark 4.1.18. We have the canonical commutative diagram

$$\begin{tikzcd}
0 \arrow[r] & \mathcal{C}_0(X, \mathbb{M}_N) \arrow[r] \arrow[d, equal] & A \arrow[r] \arrow[d, "\iota_0 \oplus \iota_1"] & B \arrow[r] \arrow[d, "\tau_0 \oplus \tau_1"'] & 0 \\
0 \arrow[r] & \bigoplus\limits_{\substack{*=0,1 \\ j \in J_*}} \mathcal{C}_0(C_j, \mathbb{M}_N) \arrow[r] & \bigoplus\limits_{*=0,1} \prod\limits_{j \in J_*} \mathcal{C}(\beta C_j, \mathbb{M}_N) \arrow[r, "q_0 \oplus q_1"] & \bigoplus\limits_{*=0,1} \frac{\prod_{j \in J_*} \mathcal{C}(\beta C_j, \mathbb{M}_N)}{\bigoplus_{j \in J_*} \mathcal{C}_0(C_j, \mathbb{M}_N)} \arrow[r] & 0
\end{tikzcd}$$

where $\tau_0 \oplus \tau_1$ is the Busby map τ and the right square is a pullback diagram. Since we can pass to any strongly unitarily equivalent extension (in the sense of [Bla06, II.8.4.12]) without changing the retract map R, we can, by Proposition 4.1.9 and the finiteness condition on R, assume that for every j the image of

$$\tau_j \colon B \xrightarrow{\tau} \frac{\prod_{j'} \mathcal{C}(\beta C_{j'}, \mathbb{M}_N)}{\bigoplus_{j'} \mathcal{C}_0(C_{j'}, \mathbb{M}_N)} \to \frac{\mathcal{C}(\beta C_j, \mathbb{M}_N)}{\mathcal{C}_0(C_j, \mathbb{M}_N)} = \mathcal{C}(\chi(C_j), \mathbb{M}_N)$$

is locally constant on $\chi(C_j)$, and even constant if $j \in J_1$. Furthermore, using lower semicontinuity of R and arguing as in the proof of Corollary 4.1.11, we may assume that

$$\iota_1(A) \subseteq \prod_{j \in J_1} \mathbb{M}_N \cdot 1_{\beta C_j} + \bigoplus_{j \in J_1} \mathcal{C}_0(C_j, \mathbb{M}_N).$$

Next, we write $\alpha X = X \cup \{\infty\}$ as a limit of finite graphs. By Theorem 2.1.5 we can find a sequence of finite graphs $X_i \subset X_{i+1} \subset \alpha X$ such that $X_0 = \mathrm{core}(\alpha X, \infty)$ (in the sense of 4.1.15) and each X_{i+1} is obtained from X_i by attaching a line segment $[0,1]$ at the 0-endpoint to a single point y_i of X_i. Furthermore we have $\varprojlim X_i = \alpha X$ along the sequence of first point maps $\varrho^i_\infty \colon \alpha X \to X_i$. We need to fix some notation: Denote the inclusion of X_i into X_{i+1} by ι_i^{i+1} and the retract from X_{i+1} to X_i by collapsing the attached interval to the attaching point y_i by ϱ^i_{i+1}. An analogous notation is used for the inclusion $X_i \subseteq \alpha X$:

$$\begin{tikzcd}
X_i \arrow[r, "\iota_i^{i+1}"'] & X_{i+1} \arrow[l, bend right, "\varrho^i_{i+1}"'] \arrow[r, "\iota_{i+1}^{\infty}"'] & \alpha X \arrow[l, bend right, "\varrho^{i+1}_\infty"']
\end{tikzcd}$$

Now for every pair of indices i, j we have $X_i \cap C_j$ sitting inside C_j. Note that $X_{i+1} \backslash X_i \cap C_{j(i)} \neq \emptyset$ for a unique $j(i) \in J$ since $\infty \in X_0$. We define suitable compactifications $\alpha_j(X_i \cap C_j)$ of $X_i \cap C_j$ as follows: if $X_0 \cap C_j = \emptyset$, we let $\alpha_j(X_i \cap C_j) = \alpha(X_i \cap C_j)$ be the usual one-point compactification for any $i \in \mathbb{N}$. In the case $X_0 \cap C_j \neq \emptyset$, which will occur only finitely many times, we have an inclusion $\mathcal{C}_b(X_i \cap C_j) \subseteq \mathcal{C}_b(C_j)$ induced by the surjective retract $\varrho^i_{\infty|C_j} \colon C_j \to X_i \cap C_j$ and we define $\alpha_j(X_i \cap C_j)$ via

$$\mathcal{C}(\alpha_j(X_i \cap C_j)) = \left\{ f \in \mathcal{C}_b(X_i \cap C_j) \subseteq \mathcal{C}_b(C_j) = \mathcal{C}(\beta C_j) \colon \begin{array}{c} f \text{ is locally} \\ \text{constant} \\ \text{on } \chi(C_j) \end{array} \right\}.$$

Since the corona space $\chi(C_j)$ has only finitely many connected components by Lemma 4.1.17, $\alpha_j(X_i \cap C_j)$ will be a finite-point compactification of $X_i \cap C_j$ (meaning that $\alpha_j(X_i \cap C_j) \backslash (X_i \cap C_j)$ is a finite set). In particular, $\alpha_j(X_i \cap C_j)$ is a finite graph for any pair of indices i and j. We are now ready to iteratively define the C^*-algebras B_i as the pullbacks over

$$\begin{tikzcd} B_i \arrow[r, dotted] \arrow[d, dotted] & B \arrow[d, "\tau"] \\ \prod_j \mathcal{C}(\alpha_j(X_i \cap C_j), \mathbb{M}_N) \arrow[r, "q"] & \frac{\prod_j \mathcal{C}(\beta C_j, \mathbb{M}_N)}{\bigoplus_j \mathcal{C}_0(C_j, \mathbb{M}_N)} \end{tikzcd}$$

with respect to the inclusions $(\varrho^i_{\infty|C_j})^* \otimes \mathrm{id}_{\mathbb{M}_N} \colon \mathcal{C}(\alpha_j(X_i \cap C_j), \mathbb{M}_N) \subseteq \mathcal{C}(\beta C_j, \mathbb{M}_N)$. Let us first simplify the description of B_i. For every fixed i, the set $X_i \cap C_j$ is empty for almost every $j \in J$ so that $\mathcal{C}(\alpha_j(X_i \cap C_j), \mathbb{M}_N) = \mathbb{M}_N \cdot 1_{\beta C_j}$ for almost every j. Given $((f_j)_j, b) \in B_i$, this implies $f_j = \tau_j(b) \cdot 1_{\beta C_j}$ for almost every j. Hence B_i is isomorphic to the pullback

$$\begin{tikzcd} B_i \arrow[r, dotted] \arrow[d, dotted] & B \arrow[d, "\bigoplus\limits_{j \in J(i)} \tau_j"] \\ \bigoplus\limits_{j \in J(i)} \mathcal{C}(\alpha_j(X_i \cap C_j), \mathbb{M}_N) \arrow[r, "q"] & \bigoplus\limits_{j \in J(i)} \frac{\mathcal{C}(\beta C_j, \mathbb{M}_N)}{\mathcal{C}_0(C_j, \mathbb{M}_N)} \end{tikzcd}$$

for the finite set $J(i) = \{j \in J \colon X_i \cap C_j \neq \emptyset\} \subseteq J$. Since every $\alpha(X_i \cap C_j)$ is a finite graph, the C^*-algebra on the lower left side is a 1-NCCW. One also checks that the pullbacks are taken over finite-dimensional C^*-algebras because $(\oplus_{j \in J(i)} \tau_j)(B)$ consists of locally constant functions on the space $\bigsqcup_{j \in J(i)} \chi(C_j)$ which has only finitely many connected components by Lemma 4.1.17.
Next, we specify the inductive structure, i.e. the connecting maps $s_i^{i+1} \colon B_i \to B_{i+1}$ and retracts $r^i_{i+1} \colon B_{i+1} \to B_i$. By definition, we find $B_i \subseteq B_{i+1} \subseteq A$ with the inclusions coming from $(\varrho^i_{i+1})^* \otimes \mathrm{id}_{\mathbb{M}_N}$ resp. by $(\varrho^{i+1}_\infty)^* \otimes \mathrm{id}_{\mathbb{M}_N}$. We denote them by s_i^{i+1} resp. by s_i^∞. Since $\overline{\bigcup_i \mathcal{C}(\alpha_j(X_i \cap C_j), \mathbb{M}_N)} \supseteq \overline{\bigcup_i \mathcal{C}_0(X_i \cap C_j, \mathbb{M}_N)} = \mathcal{C}_0(X \cap C_j, \mathbb{M}_N)$ for every $j \in J$, we find $\mathcal{C}_0(X, \mathbb{M}_N) \subseteq \overline{\bigcup_i B_i}$. One further checks that $\bigoplus_{j \in J_0} \mathcal{C}(\alpha_j(X_0 \cap C_j), \mathbb{M}_N)$ surjects via q onto the locally constant functions on $\bigsqcup_{j \in J_0} \chi(C_j)$. Together with $\tau_1(B) \subseteq q_1(\prod_{j \in J_1} \mathbb{M}_N \cdot 1_{\beta C_j}) \subseteq q_1(\prod_{j \in J_1} \mathcal{C}(\alpha_j(X_0 \cap C_j), \mathbb{M}_N))$ it follows that $\overline{\bigcup_i B_i}$ is the pullback over τ and q, and hence $\overline{\bigcup_i B_i} = A$.

It remains to verify the description of the connecting maps s_i^{i+1}. We have $X_i \cap C_j = X_{i+1} \cap C_j$ if $j \neq j(i)$ and $\alpha_j(X_i \cap C_{j(i)}) \subseteq \alpha_j(X_{i+1} \cap C_{j(i)}) \cong \alpha_j(X_i \cap C_{j(i)}) \cup_{\{y_i\}=\{0\}} [0,1]$. This means there is a pullback diagram

$$\begin{tikzcd} \mathcal{C}(\alpha_j(X_{i+1} \cap C_{j(i)}), \mathbb{M}_N) \arrow[r, dotted] \arrow[d, dotted] & \mathcal{C}([0,1], \mathbb{M}_N) \arrow[d, "\mathrm{ev}_0"] \\ \mathcal{C}(\alpha_j(X_i \cap C_{j(i)}), \mathbb{M}_N) \arrow[r, "\mathrm{ev}_{y_i}"] \arrow[u, bend left, "(\varrho_{i+1}^i)^* \otimes \mathrm{id}_{\mathbb{M}_N}"] & \mathbb{M}_N \end{tikzcd}$$

where $(\varrho_{i+1}^i)^* \otimes \mathrm{id}_{\mathbb{M}_N}$ corresponds to $f \mapsto (f, f(y_i) \otimes 1_{[0,1]})$ in the pullback picture and the downward arrow on the left side comes from the inclusion $\alpha_j(X_i \cap C_{j(i)}) \subseteq \alpha_j(X_{i+1} \cap C_{j(i)})$. This map induces a surjection $B_{i+1} \to B_i$ which will be denoted by r_{i+1}^i and gives the claimed pullback diagram.
Finally, if αX is an AR-space, the core $X_0 = \mathrm{core}(\alpha X, \infty) = [x_0, \infty]$ is nothing but an arc from some point $x_0 \in X$ to ∞. In this case the finite set $J(0)$ consists of a single element $j(0)$, namely the index corresponding to the component containing x_0. By definition, B_0 comes as a pullback

$$\begin{tikzcd} B_0 \arrow[r, dotted] \arrow[d, dotted] & B \arrow[d, "\tau_{j(0)}"] \\ \mathcal{C}([x_0, \infty], \mathbb{M}_N) \arrow[r, "\mathrm{ev}_\infty"] & \mathbb{M}_N \cdot 1_{\chi(C_{j(0)})} \end{tikzcd}$$

and hence an index shift allows us to start with $B_0 \cong B$. □

The procedure of forming extensions by C^*-algebras of the form $\mathcal{C}_0(X, \mathbb{M}_N)$ can of course be iterated. The next proposition shows that, if all the attached spaces X are one-dimensional ANRs up to compactification, the limit structures which we get from Lemma 4.2.1 for each step can be combined into a single one.

Proposition 4.2.2. *Let a short exact sequence of separable C^*-algebras $0 \to \mathcal{C}_0(X, \mathbb{M}_N) \to A \to B \to 0$ be given. Assume that αX is a one-dimensional ANR-space and that the associated set-valued retract map $R\colon \mathrm{Prim}(A) \to 2^{\mathrm{Prim}(B)}$ as in 4.1.1 is lower semicontinuous and has pointwise finite image. Further assume that there exists a direct limit structure for B*

$$\begin{tikzcd} B_0 \arrow[r, "s_0^1"'] & B_1 \arrow[r, "s_1^2"'] \arrow[l, bend right, "r_1^0"'] & B_2 \arrow[r] \arrow[l, bend right, "r_2^1"'] & \cdots \arrow[l, bend right] \arrow[r, "s_{i-1}^i"'] & B_i \arrow[r] \arrow[l, bend right, "r_i^{i-1}"'] \arrow[rr, bend right, "s_i^\infty"'] & \cdots \arrow[r] & B \arrow[ll, bend right, "r_\infty^i"'] \end{tikzcd}$$

such that all B_i are 1-NCCWs and at each stage there is a representation

$p_i\colon B_i \to \mathbb{M}_{n_i}$ such that B_{i+1} is defined as the pullback

$$\begin{tikzcd}
B_{i+1} \arrow[r, dotted, "t_{i+1}"] \arrow[d, dotted, two heads, "r^i_{i+1}"] & \mathcal{C}([0,1],\mathbb{M}_{n_i}) \arrow[d, "\mathrm{ev}_0"] \\
B_i \arrow[r, "p_i"] \arrow[u, bend left, "s_i^{i+1}"] & \mathbb{M}_{n_i}
\end{tikzcd}$$

and $s_i^{i+1}\colon B_i \to B_{i+1}$ is given by $a \mapsto (a, p_i(a) \otimes 1_{[0,1]})$.
Then A is isomorphic to the limit A_∞ of an inductive system

$$\begin{tikzcd}
A_0 \arrow[r, "\sigma_0^1"'] & A_1 \arrow[l, bend right, "\varrho_1^0"'] \arrow[r, "\sigma_1^2"'] & A_2 \arrow[l, bend right, "\varrho_2^1"'] \arrow[r] & \cdots \arrow[l, bend right] \arrow[r, "\sigma_{i-1}^i"'] & A_i \arrow[l, bend right, "\varrho_i^{i-1}"'] \arrow[r] \arrow[rr, bend right, "\sigma_i^\infty"'] & \cdots \arrow[r] & A_\infty \arrow[ll, bend right, "\varrho_\infty^i"']
\end{tikzcd}$$

where all A_i are 1-NCCWs and at each stage there is a representation $\pi_i\colon A_i \to \mathbb{M}_{m_i}$ such that A_{i+1} is defined as the pullback

$$\begin{tikzcd}
A_{i+1} \arrow[r, dotted] \arrow[d, dotted, two heads, "\varrho^i_{i+1}"] & \mathcal{C}([0,1],\mathbb{M}_{m_i}) \arrow[d, "\mathrm{ev}_0"] \\
A_i \arrow[r, "\pi_i"] \arrow[u, bend left, "\sigma_i^{i+1}"] & \mathbb{M}_{m_i}
\end{tikzcd}$$

and $\sigma_i^{i+1}\colon A_i \to A_{i+1}$ is given by $a \mapsto (a, \pi_i(a) \otimes 1_{[0,1]})$. Furthermore, if αX is an AR-space we may even arrange $A_0 \cong B_0$.

Proof. By Lemma 4.2.1, we know that A can be written as an inductive limit

$$\begin{tikzcd}
\overline{A}_0 \arrow[r, "\overline{s}_0^1"'] & \overline{A}_1 \arrow[l, bend right, "\overline{r}_1^0"'] \arrow[r, "\overline{s}_1^2"'] & \overline{A}_2 \arrow[l, bend right, "\overline{r}_2^1"'] \arrow[r] & \cdots \arrow[l, bend right] \arrow[r, "\overline{s}_{i-1}^i"'] & \overline{A}_i \arrow[l, bend right, "\overline{r}_i^{i-1}"'] \arrow[r] \arrow[rr, bend right, "\overline{s}_i^\infty"'] & \cdots \arrow[r] & A \arrow[ll, bend right, "\overline{r}_\infty^i"']
\end{tikzcd}$$

with a pullback structure

$$\begin{tikzcd}
\overline{A}_{i+1} \arrow[r, dotted] \arrow[d, dotted, two heads, "\overline{r}^i_{i+1}"] & \mathcal{C}([0,1],\mathbb{M}_N) \arrow[d, "\mathrm{ev}_0"] \\
\overline{A}_i \arrow[r, "\overline{p}_i"] \arrow[u, bend left, "\overline{s}_i^{i+1}"] & \mathbb{M}_N
\end{tikzcd}$$

at every stage and $\overline{s}_i^{i+1}\colon \overline{A}_i \to \overline{A}_{i+1}$ given by $a \mapsto (a, \overline{p}_i(a) \otimes 1_{[0,1]})$. The starting algebra $\overline{A}_0$ comes as a pullback

$$\begin{tikzcd}
\overline{A}_0 \arrow[r] \arrow[d] & D \arrow[d, "\varphi"] \\
B \arrow[r, "\psi"] & F
\end{tikzcd}$$

with D a 1-NCCW and $\dim(F) < \infty$. In the case of αX being an AR-space, we may choose $\overline{A}_0 = B$, i.e. $D = 0$. For $j \in \mathbb{N}$ we now define $A_{0,j}$ to be the pullback

$$\begin{tikzcd}
A_{0,j} \arrow[r, dotted] \arrow[d, dotted, "\varrho_{0,j}"'] & D \arrow[d, "\varphi"] \\
B_j \arrow[r, "\psi\circ s_j^{\infty}"] & F.
\end{tikzcd}$$

The maps s_j^{j+1},s_j^{∞} homomorphisms $\sigma_{0,j}^{0,j+1}\colon A_{0,j} \to A_{0,j+1}$ and $\sigma_{0,j}^{0,\infty}\colon A_{0,j} \to \overline{A}_0$, leading to an inductive limit structure with $\varinjlim_j(A_{0,j}, \sigma_{0,j}^{0,j+1}) = \overline{A}_0$. We proceed iteratively, defining $A_{i+1,j}$ to be the pullback

$$\begin{tikzcd}
A_{i+1,j} \arrow[r, dotted] \arrow[d, dotted, "\varrho_{i+1,j}^{i,j}"] & \mathcal{C}([0,1], \mathbb{M}_N) \arrow[d, "\mathrm{ev}_0"] \\
A_{i,j} \arrow[u, dotted, bend left, "\sigma_{i,j}^{i+1,j}"] \arrow[r, "\overline{p}_i\circ\sigma_{i,j}^{i,\infty}"'] & \mathbb{M}_N
\end{tikzcd}$$

with $\sigma_{i,j}^{i+1,j}\colon A_{i,j} \to A_{i+1,j}$ given by $a \mapsto (a, (\overline{p}_i \circ \sigma_{i,j}^{i,\infty})(a) \otimes 1_{[0,1]})$. It is then checked that $\sigma_{i,j}^{i,j+1}$ and $\sigma_{i,j}^{i,\infty}$ induce compatible homomorphisms $\sigma_{i+1,j}^{i+1,j+1}\colon A_{i+1,j} \to A_{i+1,j+1}$ and $\sigma_{i+1,j}^{i+1,\infty}\colon A_{i+1,j} \to \overline{A}_{i+1}$ with $\varinjlim_j(A_{i+1,j}, \sigma_{i+1,j}^{i+1,j+1}) = \overline{A}_{i+1}$. Observing that for every i and j

$$\begin{tikzcd}
A_{i,j+1} \arrow[r, "t_{j+1}\circ\varrho_{0,j+1}\circ\varrho_{i,j+1}^{0,j+1}"] \arrow[d, "\varrho_{i,j+1}^{i,j}"] & \mathcal{C}([0,1], \mathbb{M}_{n_j}) \arrow[d, "\mathrm{ev}_0"] \\
A_{i,j} \arrow[u, bend left, "\sigma_{i,j}^{i,j+1}"] \arrow[r, "p_j\circ\varrho_{0,j}\circ\varrho_{i,j}^{0,j}"'] & \mathbb{M}_{n_j}
\end{tikzcd}$$

is indeed a pullback diagram, we get the desired limit structure for A by following the diagonal in the commutative diagram

$$\begin{tikzcd}
A_{00} \arrow[r] \arrow[d, dotted] & A_{01} \arrow[r] \arrow[d] & A_{02} \arrow[r] \arrow[d] & A_{03} \arrow[r] \arrow[d] & \cdots \arrow[r] & \overline{A}_0 \arrow[d] \\
A_{10} \arrow[r, dotted] \arrow[d] & A_{11} \arrow[r] \arrow[d, dotted] & A_{12} \arrow[r] \arrow[d] & A_{13} \arrow[r] \arrow[d] & \cdots \arrow[r] & \overline{A}_1 \arrow[d] \\
A_{20} \arrow[r] \arrow[d] & A_{21} \arrow[r, dotted] \arrow[d] & A_{22} \arrow[r] \arrow[d, dotted] & A_{23} \arrow[r] \arrow[d] & \cdots \arrow[r] & \overline{A}_2 \arrow[d] \\
A_{30} \arrow[r] \arrow[d] & A_{31} \arrow[r] \arrow[d] & A_{32} \arrow[r, dotted] \arrow[d] & A_{33} \arrow[r] \arrow[d, dotted] & \cdots \arrow[r] & \overline{A}_3 \arrow[d] \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots
\end{tikzcd}$$

as indicated. Note that, since all connecting maps are injective, the limit over the diagonal equals $\varinjlim \overline{A}_n = A$. $\square$

4.3 Keeping track of semiprojectivity

We now reap the fruit of our labor in the previous sections and work out a '2 out of 3'-type statement describing the behavior of semiprojectivity with respect to extensions by homogeneous C^*-algebras. While for general extensions the behavior of semiprojectivity is either not at all understood or known to be rather bad, Theorem 4.3.2 gives a complete and satisfying answer in the case of homogeneous ideals. It is the very first result of this type and allows to understand semiprojectivity for C^*-algebras which are built up from homogeneous pieces, see chapter 5.1.

Proposition 4.3.1. *Let $0 \to \mathcal{C}_0(X, \mathbb{M}_N) \to A \to B \to 0$ be a short exact sequence of C^*-algebras. If both A and B are (semi)projective, then the one-point compactification αX is a one-dimensional A(N)R-space.*

Proof. The projective case follows directly from Corollary 3.1.3 and Theorem 2.3.7 while the semiprojective case requires some more work. By Lemma 2.3.8 we know that αX is a Peano space of dimension at most 1. The proof of 2.3.8 further shows that there are no small circles accumulating in X. However, in order to show that αX is an ANR-space we have to exclude the possibility of smaller and smaller circles accumulating at $\infty \in \alpha X$, see Theorem 2.1.7. Assume that we find a sequence of circles with diameters converging to 0 (with respect to some fixed geodesic metric ([Bin49], [Moi49])) at around $\infty \in \alpha X$. After passing to a subsequence, there are two possible situations: either each circle contains the point ∞ or none of them does. Both cases are treated exactly the same, for the sake of brevity we only consider the situation where ∞ is contained in all circles. In this case have a *-homomorphism $\varphi\colon \mathcal{C}_0(X,\mathbb{M}_N) \to \bigoplus_{n=1}^{\infty} \mathcal{C}_0((0,1)_n,\mathbb{M}_N)$ where $(0,1)_n \cong (0,1)$ is the part of the n-th circle contained in X. Note that each coordinate projection gives a surjection $\varphi_n\colon \mathcal{C}_0(X,\mathbb{M}_N) \to \mathcal{C}_0((0,1),\mathbb{M}_N)$ while φ itself is not necessarily surjective (because the circles might intersect in X). We make use of Brown's mapping telescope associated to $\bigoplus_{n=1}^{\infty} \mathcal{C}_0((0,1)_n,\mathbb{M}_N)$, i.e.

$$T_k = \left\{ f \in \mathcal{C}([k,\infty], \bigoplus_{n=1}^{\infty} \mathcal{C}_0((0,1)_n,\mathbb{M}_N))\colon t \leq l \Rightarrow f(t) \in \bigoplus_{n=1}^{l} \mathcal{C}_0((0,1)_n,\mathbb{M}_N) \right\}$$

with the obvious (surjective) restrictions as connecting maps giving $\varinjlim T_k = \bigoplus_{n=1}^{\infty} \mathcal{C}_0((0,1)_n,\mathbb{M}_N)$. Using Lemma 3.1.4, we find a commutative diagram

$$\begin{tikzcd}
0 \arrow[r] & \varinjlim T_k \arrow[r] & \varinjlim \mathcal{M}(T_k) \arrow[r] & \varinjlim \mathcal{Q}(T_k) \arrow[r] & 0 \,. \\
0 \arrow[r] & \mathcal{C}_0(X,\mathbb{M}_N) \arrow[r] \arrow[u, "\varphi"] & A \arrow[r] \arrow[u, "\overline{\varphi}"] & B \arrow[r] \arrow[u, "\overline{\overline{\varphi}}"] & 0
\end{tikzcd}$$

It now follows from the semiprojectivity assumptions and Lemma 3.1.1 that φ lifts to T_k for some k, which is equivalent to a solution of the original lifting problem

$$\begin{tikzcd}
 & \bigoplus\limits_{n=1}^{k} \mathcal{C}_0((0,1)_n,\mathbb{M}_N) \arrow[d, "\subseteq"] \\
\mathcal{C}_0(X,\mathbb{M}_N) \arrow[r, "\varphi"] \arrow[ur, dotted] & \bigoplus\limits_{n=1}^{\infty} \mathcal{C}_0((0,1)_n,\mathbb{M}_N)
\end{tikzcd}$$

up to homotopy. This, however, implies

$$\begin{aligned}
\mathrm{im}(K_1(\varphi)) \subseteq \quad & K_1\left(\bigoplus_{n=1}^{k} \mathcal{C}_0((0,1)_n,\mathbb{M}_N)\right) \\
= \quad & \sum_{n=1}^{k} \mathbb{Z} \\
\subset \quad & \sum_{n=1}^{\infty} \mathbb{Z} = K_1\left(\bigoplus_{n=1}^{\infty} \mathcal{C}_0((0,1)_n,\mathbb{M}_N)\right)
\end{aligned}$$

which gives a contradiction as follows. Because φ_{k+1} is surjective and $\dim(\alpha X) \leq 1$, we can extend the canonical unitary function from $\alpha((0,1)_n)$ to a unitary u on all of αX by [HW48, Theorem VI.4]. This unitary then satisfies $u - 1 \in \mathcal{C}_0(X)$ and $K_1(\varphi)([u \otimes 1_{\mathbb{M}_N}]) = N \in \mathbb{Z} = K_1(\mathcal{C}_0((0,1)_{k+1},\mathbb{M}_N))$. This shows that there are no small circles at around ∞ in αX and hence that αX is a one-dimensional ANR-space by Theorem 2.1.7. $\square$

THEOREM 4.3.2. *Let a short exact sequence of C^*-algebras $0 \to I \to A \to B \to 0$ be given and assume that the ideal I is a N-homogeneous C^*-algebra with $\mathrm{Prim}(I) = X$. Denote by $(X_i)_{i\in I}$ the connected components of X and consider the following statements:*

(I) I is (semi)projective.

(II) A is (semi)projective.

(III) B is (semi)projective and the set-valued retract map $R\colon \mathrm{Prim}(A) \to 2^{\mathrm{Prim}(B)}$ given as in 4.1.1 by

$$R(z) = \begin{cases} z & \text{if } z \in \mathrm{Prim}(B), \\ \partial X_i = \overline{X_i} \backslash X_i & \text{if } z \in X_i \subseteq X = \mathrm{Prim}(I) \end{cases}$$

is lower semicontinuous and has pointwise finite image.

If any two of these statements are true, then the third one also holds.

Proof. (I)+(II)⇒(III): By Theorem 2.3.11, we know that the sequence is isomorphic to an extension

$$0 \longrightarrow \mathcal{C}_0(X, \mathbb{M}_N) \longrightarrow A \xrightarrow{\pi} B \longrightarrow 0$$

with the one-point compactification of X a one-dimensional A(N)R-space. The set-valued retract map R is then lower semicontinuous by Proposition 4.1.20 and has pointwise finite image by Proposition 4.1.19. But now Theorem 4.1.14 applies and shows that there is a completely positive split s for the quotient map π such that the composition $B \xrightarrow{s} A \xrightarrow{\iota} \mathcal{C}_b(X, \mathbb{M}_N)$ is multiplicative outside of an open set $U \subset K \subset X$ where K is compact.

Let a lifting problem $\varphi\colon B \xrightarrow{\sim} D/J = \varinjlim D/J_n$ be given. Since A is semiprojective, we can solve the resulting lifting problem for A, meaning we find $\psi\colon A \to D/J_n$ for some n with $\pi_n \circ \psi = \varphi \circ \pi$. Restricting to $\mathrm{her}_{D/J_n}(\psi(\mathcal{C}_0(X, \mathbb{M}_N))) + \psi(A) \subseteq D/J_n$, we may assume that $\psi_{|\mathcal{C}_0(X,\mathbb{M}_N)}$ is proper as a *-homomorphism to J/J_n and hence induces a map $\mathcal{M}(\psi)$ between multiplier algebras. Since the restriction of $\pi_n \circ \psi$ to the ideal $\mathcal{C}_0(X, \mathbb{M}_N)$ vanishes, we may use compactness of K to assume that ψ maps $\mathcal{C}_0(U, \mathbb{M}_N)$ to 0 (after increasing n if necessary). This further implies that $\mathcal{M}(\psi)$ factors through $r\colon \mathcal{C}_b(X, \mathbb{M}_N) \to \mathcal{C}_b(X\backslash U, \mathbb{M}_N)$. We then find $s' := r \circ \iota \circ s$ to be multiplicative and hence a *-homomorphism:

$$\begin{tikzcd}
A \arrow[rrr, "\psi"] \arrow[dr, "\iota"] \arrow[ddd, "\pi"] & & & D/J_n \arrow[dr, "\iota_n"] \arrow[ddd, two heads, "\pi_n"] & \\
& \mathcal{C}_b(X,\mathbb{M}_N) \arrow[rrr, "\mathcal{M}(\psi)"] \arrow[dr, "r"] & & & \mathcal{M}(J/J_n) \arrow[dd, "\varrho_n"] \\
& & \mathcal{C}_b(X\backslash U,\mathbb{M}_N) \arrow[urr, dashed, "\mathcal{M}(\psi)'"] & & \\
B \arrow[rrr, "\varphi"] \arrow[urr, dashed, "s'"] \arrow[uuu, bend left, "s"] & & & D/J \arrow[r, "\tau_n"] & \mathcal{Q}(J/J_n)
\end{tikzcd}$$

The inclusion of J/J_n as an ideal in D/J_n gives canonical *-homomorphisms ι_n and τ_n as in the diagram above. One now checks that $\varrho_n\circ(\mathcal{M}(\psi)'\circ s')=\tau_n\circ\varphi$ holds. Combining this with the fact that the trapezoid on the right is a pullback diagram, we see that the pair $(\varphi,(\mathcal{M}(\psi)'\circ s'))$ defines a lift $B\to D/J_n$ for φ. This shows that the quotient B is semiprojective.
For the projective version of the statement, one uses Corollary 4.1.11 to see that the sequence admits a multiplicative split $s\colon B\to A$ rather than just a completely positive one.

(I)+(III)⇒(II): We know that $I\cong\mathcal{C}_0(X,\mathbb{M}_N)$ with αX a one-dimensional A(N)R-space by Theorem 2.3.11. Now Lemma 4.2.1 applies and we obtain a limit structure for A

$$\begin{tikzcd}
B_0 \arrow[r, "s_0^1"'] & B_1 \arrow[l, bend right, "r_1^0"'] \arrow[r, "s_1^2"'] & B_2 \arrow[l, bend right, "r_2^1"'] \arrow[r] & \cdots \arrow[l, bend right] \arrow[r, "s_{i-1}^i"'] & B_i \arrow[l, bend right, "r_i^{i-1}"'] \arrow[r] \arrow[rr, bend right, "s_i^\infty"'] & \cdots \arrow[r] & \varinjlim\left(B_i, s_i^{i+1}\right)\cong A \arrow[ll, bend right, "r_\infty^i"']
\end{tikzcd}$$

with B_0 given as a pullback of B and a 1-NCCW D over a finite-dimensional C^*-algebra. In particular, B_0 is semiprojective by [End14, Corollary 3.4]. In the projective case, we can take $B_0=B$ to be projective. In both cases, the connecting maps in the system above arise from pullback diagrams

$$\begin{tikzcd}
B_{i+1} \arrow[r, dotted] \arrow[d, dotted, two heads, "r_{i+1}^i"] & \mathcal{C}([0,1],\mathbb{M}_N) \arrow[d, "\mathrm{ev}_0"] \\
B_i \arrow[u, bend left, "s_i^{i+1}"] \arrow[r, "\pi_i"] & \mathbb{M}_N
\end{tikzcd}$$

with $s_i^{i+1}(a)=(a,\pi_i(a)\otimes 1_{[0,1]})$. Since these maps are weakly conditionally projective by Proposition 3.2.4, we obtain (semi)projectivity of A from Lemma 3.2.3.

(II)+(III)⇒(I): This implication holds under even weaker hypothesis. More precisely, we show that (semi)projectivity of both A and B implies I to be (semi)projective. The assumption on the retract map R is not needed here. First we apply Lemma 2.3.8 to find the one-point compactification of $\mathrm{Prim}(I)$ to be a Peano space of dimension at most 1, and hence I is trivially homogeneous by Lemma 2.3.10. Now Proposition 4.3.1 shows that αX is in fact an ANR-space which, together with Theorem 2.3.11, means that I is semiprojective. The projective version is Corollary 3.1.3. □

Remark 4.3.3. *Theorem 4.3.2 shows that regularity properties of the retract map* $R\colon \mathrm{Prim}(A) \to 2^{\mathrm{Prim}(B)}$ *are crucial for semiprojectivity to behave nicely with respect to extensions by homogeneous* C^**-algebras. This can already be observed and illustrated in the commutative case. Given an extension of commutative* C^**-algebras*

$$0 \to \mathcal{C}_0(X) \to \mathcal{C}_0(Y) \to \mathcal{C}_0(Y \backslash X) \to 0,$$

the following holds: If both the ideal $\mathcal{C}_0(X)$ *and the quotient* $\mathcal{C}_0(Y\backslash X)$ *are (semi)projective, then the extension* $\mathcal{C}_0(Y)$ *is (semi)projective if and only if the retract map* $R\colon Y \to 2^{Y\backslash X}$ *is lower semicontinuous and has pointwise finite image. The following examples show that both properties for* R *are not automatic:*

(a) An examples with R *not having pointwise finite image is contained as example 5.5 in [LP98], we include it here for completeness. Let* $X = \{(x, \sin x^{-1}) : 0 < x \leq 1\} \subset \mathbb{R}^2$ *and* $Y = X \cup \{(0, y)\colon -1 \leq y < 1\}$*, then we get an extension isomorphic to*

$$0 \to \mathcal{C}_0(0,1] \to \mathcal{C}_0(Y) \to \mathcal{C}_0(0,1] \to 0.$$

Here both the ideal and the quotient are projective, but the extension $\mathcal{C}_0(Y)$ *is not (because* αY *is not locally connected and hence not an AR-space). In this example, we find* $R(x) = \{(0, y)\colon -1 \leq y < 1\}$ *to be infinite for all* $x \in X$*.*

(b) The following is an example where R *fails to be lower semicontinuous. Consider* $Y = \{(x, 0)\colon 0 \leq x < 1\} \cup \bigcup_n C_n \subset \mathbb{R}^2$ *with* $C_n = \{(t, (1-t)/n)\colon 0 \leq t < 1\}$ *the straight line from* $(0, 1/n)$ *to* $(1, 0)$ *with the endpoint* $(1, 0)$ *removed. With* $X = \bigcup_n C_n \subset Y$ *we obtain an extension isomorphic to*

$$0 \to \bigoplus_n \mathcal{C}_0(0,1] \to \mathcal{C}_0(Y) \to \mathcal{C}_0(0,1] \to 0.$$

Here both the ideal and the quotient are projective while the extension $\mathcal{C}_0(Y)$ *is not (again because* αY *is not locally connected). We also find* $(0, 1/n) \to (0, 0)$ *in* Y *but* $R((0, 1/n)) = \emptyset$ *for all* n*, which shows that* R *is not lower semicontinuous. The descriptive reason for* $\mathcal{C}_0(Y)$ *not being projective in this case is that the length of the attached intervals* C_n *does not tend to* 0 *as* n *goes to infinity.*

5 The structure of semiprojective subhomogeneous C^*-algebras

5.1 The main result

With Theorem 4.3.2 at hand, we are now able to keep track of semiprojectivity when decomposing a subhomogeneous C^*-algebra into its homogeneous subquotients. On the other hand, Theorem 4.3.2 also tells us in which manner homogeneous, semiprojective C^*-algebras may be combined in order to give subhomogeneous C^*-algebras which are again semiprojective. This leads to the main result of this chapter, Theorem 5.1.2, which gives two characterizations of projectivity and semiprojectivity for subhomogeneous C^*-algebras.

Lemma 5.1.1. *Let A be a N-subhomogeneous C^*-algebra. If A is semiprojective, then the maximal N-homogeneous ideal of A is also semiprojective.*

Proof. By Lemma 2.3.8 we know that the one-point compactification of $X = \mathrm{Prim}_N(A)$ is a one-dimensional Peano space. Since any locally trivial $\mathbb{M}_N$-bundle over X is globally trivial by Lemma 2.3.10, we are concerned with an extension of the form

$$0 \longrightarrow \mathcal{C}_0(X,\mathbb{M}_N) \longrightarrow A \xrightarrow{\pi} A_{\leq N-1} \longrightarrow 0$$

where $A_{\leq N-1}$ denotes the maximal (N-1)-subhomogeneous quotient of A. Since A is semiprojective, $A_{\leq N-1}$ will be semiprojective with respect to (N-1)-subhomogeneous C^*-algebras. In order to show that $\mathcal{C}_0(X,\mathbb{M}_N)$ is semiprojective, it remains to show that $\alpha X = X \cup \{\infty\}$ does not contain small circles at around ∞, cf. Theorem 2.1.7. The proof for this is similar to the one of 4.3.1. We use notations from 2.3.8 and follow the proof there to arrive at a commutative diagram

$$\begin{tikzcd}
0 \arrow[r] & \varinjlim T_k \arrow[r] & \varinjlim \mathcal{M}(T_k) \arrow[r] & \varinjlim \mathcal{Q}(T_k) \arrow[r] & 0\ . \\
0 \arrow[r] & \mathcal{C}_0(X,\mathbb{M}_N) \arrow[r] \arrow[u, "\varphi"] & A \arrow[r, "\pi"] \arrow[u, "\overline{\varphi}"] & A_{\leq N-1} \arrow[r] \arrow[u, "\overline{\overline{\varphi}}"] & 0
\end{tikzcd}$$

We may not solve the lifting problem for $A_{\leq N-1}$ directly since the algebras $\mathcal{Q}(T_k)$ are not (N-1)-subhomogeneous. Instead we will replace the $\mathcal{Q}(T_k)$ by suitable (N-1)-subhomogeneous subalgebras which will then lead to a solvable lifting problem for $A_{\leq N-1}$. Let ι_n denote the n-th coordinate of the map $A \to \mathcal{C}_b(X,\mathbb{M}_N) \to \prod_n \mathcal{C}_b((0,1)_n,\mathbb{M}_N)$. We then have a lift of $\overline{\varphi}$ given by

$$\begin{array}{ccccc} A & \to & \mathcal{C}([k,\infty],\prod_n \mathcal{C}_b((0,1),\mathbb{M}_N)) & \to & \mathcal{M}(T_k) \\ a & \mapsto & 1_{[k,\infty]} \otimes (\iota_n(a))_{n=1}^{\infty} & & \end{array}$$

where the map on the right is induced by the inclusion of T_k as an ideal in $\mathcal{C}([k,\infty],\prod_n \mathcal{C}_b((0,1),\mathbb{M}_N))$. Consider in there the central element $f = (f_n)_{n=1}^{\infty}$

with f_n the scalar function that equals 0 on $[k,n]$, 1 on $[n+1,\infty]$ and which is linear in between. Then

$$\begin{array}{rcccc} \psi\colon A & \to & \mathcal{C}([k,\infty],\prod_n \mathcal{C}_b((0,1),\mathbb{M}_N)) & \to & \mathcal{M}(T_k) \\ a & \mapsto & (f_n\otimes\iota_n(a))_{n=1}^{\infty} & & \end{array}$$

is a completely positive lift of $\overline{\varphi}$ which sends $\mathcal{C}_0(X,\mathbb{M}_N)$ to T_k. Hence ψ induces a completely positive lift $\psi'\colon A_{\leq N-1}\to\mathcal{Q}(T_k)$ of $\overline{\overline{\varphi}}$. We claim that $C^*(\psi'(A_{\leq N-1}))$ is in fact (N-1)-subhomogeneous. To see this, we use the algebraic characterization of subhomogenity as described in [Bla06, IV.1.4.6]. It suffices to check that $\gamma(C^*(\psi'(A_{\leq N-1})))$ satisfies the polynomial relations $p_{r(N-1)}$ for every irreducible representation γ of $\mathcal{Q}(T_k)$. By definition of ψ, we find $\gamma\circ\psi'(\pi(a))=t\cdot\gamma'(\iota(a))$ for some representation γ' of $\iota(A)$, some $t\in[0,1]$ and every $a\in A$. Moreover, since ψ' maps $\mathcal{C}_0(X,\mathbb{M}_N)$ to 0, we obtain $\gamma\circ\psi'(\pi(a))=t\cdot\gamma''(\pi(a))$ for some representation γ'' of $A_{\leq N-1}$. Using (N-1)-subhomogeneity of $A_{\leq N-1}$, it now follows easily that the elements of $\gamma(C^*(\psi'(A_{\leq N-1})))$ satisfy the polynomial relations $p_{r(N-1)}$ from [Bla06, IV.1.4.6]. Knowing that the image of $\overline{\overline{\varphi}}$ has a (N-1)-subhomogenous preimage in $\mathcal{Q}(T_k)$, we may now solve the lifting problem for $A_{\leq N-1}$. It then follows from Lemma 3.1.1 (and its proof) that φ lifts to T_k for some k. The remainder of the proof is exactly the same as the one of Proposition 4.3.1. □

We now present two characterizations of projectivity and semiprojectivity for subhomogeneous C^*-algebras. The first one describes semiprojectivity of these algebras in terms of their primitive ideal spaces. The second description characterizes them as those C^*-algebras which arise from 1-NCCWs by adding a sequence of non-commutative edges (of bounded dimension), cf. section 3.2.2.

Theorem 5.1.2. *Let A be a N-subhomogeneous C^*-algebra, then the following are equivalent:*

(1) A is semiprojective (resp. projective).

(2) For every $n=1,...,N$ the following holds:

- *The one-point compactification of $\mathrm{Prim}_n(A)$ is an ANR-space (resp. an AR-space) of dimension at most 1.*
- *If $(X_i)_{i\in I}$ denotes the family of connected components of $\mathrm{Prim}_n(A)$, then the set-valued retract map*

$$R_n\colon\ \mathrm{Prim}_{\leq n}(A)\to 2^{\mathrm{Prim}_{\leq n-1}(A)}$$

given by

$$z \quad\mapsto \begin{cases} z & \text{if } z\in\mathrm{Prim}_{\leq n-1}(A) \\ \partial X_i & \text{if } z\in X_i\subset\mathrm{Prim}_n(A)\end{cases}$$

is lower semicontinuous and has pointwise finite image.

(3) A is isomorphic to the direct limit $\varinjlim_k (A_k, s_k^{k+1})$ of a sequence of 1-NCCWs

$$\begin{tikzcd} \cdots \arrow[r] & A_k \arrow[r, "s_k^{k+1}"'] \arrow[l, bend right] & A_{k+1} \arrow[r] \arrow[l, bend right, "r_{k+1}^k"'] & \cdots \arrow[l, bend right] \end{tikzcd}$$

(with $A_0 = 0$) such that for each stage there is a pullback diagram

$$\begin{tikzcd} A_{k+1} \arrow[r, dotted] \arrow[d, dotted, two heads, "r_{k+1}^k"] & \mathcal{C}([0,1], \mathbb{M}_n) \arrow[d, "\mathrm{ev}_0"] \\ A_k \arrow[r, "\pi_k"] \arrow[u, bend left, "s_k^{k+1}"] & \mathbb{M}_n \end{tikzcd}$$

with $n \le N$ and s_k^{k+1} given by $a \mapsto (a, \pi_k(a) \otimes 1_{[0,1]})$.

Proof. (1) $\Rightarrow$ (2): We prove the implication by induction over N. The base case $N = 1$ is given by Theorem 2.3.7. Now given a N-subhomogeneous, (semi)projective C^*-algebra A, we know by Lemma 5.1.1 that the maximal N-homogeneous ideal A_N of A is (semi)projective as well. This forces $\alpha \operatorname{Prim}_N(A)$ to be a one-dimensional A(N)R-space by Theorem 2.3.11. Applying Theorem 4.3.2 to the sequence

$$0 \to A_N \to A \to A_{\le N-1} \to 0$$

now shows that the retract map $R_N \colon \operatorname{Prim}_N(A) \to 2^{\operatorname{Prim}_{\le N-1}(A)}$ is lower semicontinuous, has pointwise finite image and that the maximal (N-1)-subhomogeneous quotient $A_{\le N-1}$ is (semi)projective. The remaining statements follow from the induction hypothesis applied to $A_{\le N-1}$.
(2) $\Rightarrow$ (3): By Lemma 2.3.10, we know that the maximal N-homogeneous ideal A_N of A is of the form $\mathcal{C}_0(\operatorname{Prim}_N(A), \mathbb{M}_N)$. Using induction over N, the statement then follows from Proposition 4.2.2 applied to the sequence

$$0 \to \mathcal{C}_0(\operatorname{Prim}_N(A), \mathbb{M}_N) \to A \to A_{\le N-1} \to 0.$$

The base case $N = 1$ is given by Theorem 2.1.5.
(3) $\Rightarrow$ (1): Note that the connecting maps are weakly conditionally projective by Proposition 3.2.4, then apply Lemma 3.2.3. $\square$

Remark 5.1.3. *The most prominent examples of subhomogeneous, semiprojective C^*-algebras are the one-dimensional non-commutative CW-complexes (1-NCCWs, see Example 2.2.3). The structure theorem 5.1.2 shows that these indeed play a special role in the class of all subhomogeneous, semiprojective C^*-algebras. By part (2) of 5.1.2, they are precisely those subhomogeneous, semiprojective C^*-algebras for which the spaces $\alpha \operatorname{Prim}_n$ are all finite graphs rather than general one-dimensional ANR-spaces. Hence 1-NCCWs should be thought of as the elements of 'finite type' in the class of subhomogeneous,* *semiprojective C^*-algebras. Moreover, part (3) of 5.1.2 shows that every subhomogeneous, semiprojective C^*-algebra can be constructed from 1-NCCWs in a very controlled manner. Therefore these algebras share many properties with 1-NCCWs, as we will see in section 5.2.1 in more detail.*

5.2 Applications

Now we discuss some consequences of Theorem 5.1.2. First we collect some properties of semiprojective, subhomogeneous C^*-algebras which follow from the descriptions in 5.1.2. This includes information about their dimension and K-theory as well as details about their relation to 1-NCCWs and some further closure properties.
At least in principle one can use the structure theorem 5.1.2 to test any given subhomogeneous C^*-algebra A for (semi)projectivity. Since this would require a complete computation of the primitive ideal space of A, it is not recommended though. Instead one might use 5.1.2 as a tool to disprove semiprojectivity for a candidate A. In fact, showing directly that a C^*-algebra A is not semiprojective can be surprisingly difficult. One might therefore take one of the conditions from 5.1.2 which are easier to verify and test A for those instead. We illustrate this strategy in section 5.2.2 by proving the quantum permutation algebras to be not semiprojective. This corrects a claim in [Bla04] on semiprojectivity of universal C^*-algebras generated by finitely many projections with order and orthogonality relations.

5.2.1 Further structural properties

By part (3) of Theorem 5.1.2, we know that any semiprojective, subhomogeneous C^*-algebra comes as a direct limit of 1-NCCWs. Since the connecting maps are explicitly given and of a very special nature, it is possible to show that these limits are approximated by 1-NCCWs in a very strong sense. The following corollary makes this approximation precise.

Corollary 5.2.1 (Approximation by 1-NCCWs). *Let A be a subhomogeneous C^*-algebra. If A is semiprojective, then for every finite set $\mathcal{G} \subset A$ and every $\epsilon > 0$ there exist a 1-NCCW $B \subseteq A$ and a *-homomorphism $r\colon A \to B$ such that $\mathcal{G} \subset_\epsilon B$ and r is a strong deformation retract for B, meaning that there exists a homotopy H_t from $H_0 = \mathrm{id}_A$ to $H_1 = r$ with $H_{t|B} = \mathrm{id}_B$ for all t. In particular, A is homotopy equivalent to a one-dimensional non-commutative CW-complex.*

Proof. Use part (3) of Theorem 5.1.2 to write $A = \varinjlim A_n$ and find a suitable 1-NCCW $B = A_{n_0}$ which almost contains the given finite set $\mathcal{G}$. It is straightforward to check that the strong deformation retracts $r_n^{n_0}\colon A_n \to A_{n_0}$ give rise to a strong deformation retract $r\colon \varinjlim A_n \to A_{n_0}$. □

In particular, 1-NCCWs and semiprojective, subhomogeneous C^*-algebras share the same homotopy invariant properties. For example, we obtain the following restrictions on the K-theory of these algebras:

COROLLARY 5.2.2. *Let A be a subhomogeneous C^*-algebra. If A is semiprojective, then its K-theory is finitely generated and $K_1(A)$ is torsion free.*

Another typical phenomenon of (nuclear) semiprojective C^*-algebras is that they appear to be one-dimensional in some sense. In the context of subhomogeneous C^*-algebras, we can now make this precise using the notion of topological dimension which, for subhomogeneous A, is given by $\operatorname{topdim}(A) = \max_n \dim(\alpha \operatorname{Prim}_n(A))$.

COROLLARY 5.2.3. *Let A be a subhomogeneous C^*-algebra. If A is semiprojective, then A has stable rank 1 and* $\operatorname{topdim}(A) \leq 1$.

Proof. The statement on the stable rank of A follows from Corollary 5.2.1, while the topological dimension can be estimated using part (2) of Theorem 5.1.2. □

Our structure theorem can also be used to study permanence properties of semiprojectivity when restricted to the class of subhomogeneous C^*-algebras. In fact, these turn out to be way better then in the general situation. This can be illustrated by the following longstanding question by Blackadar and Loring: Given a short exact sequence of C^*-algebras

$$0 \longrightarrow I \longrightarrow A \longrightarrow F \longrightarrow 0$$

with finite-dimensional F, does the following hold?

$$I \text{ semiprojective} \Leftrightarrow A \text{ semiprojective}$$

While we showed the '$\Leftarrow$'-implication to hold in general in [End14], S. Eilers and T. Katsura proved the '$\Rightarrow$'-implication to be wrong ([EK]), even in the case of split extensions by $\mathbb{C}$. We refer the reader to [Sør12] for counterexamples which involve infinite C^*-algebras. However, when one restricts to the class of subhomogeneous C^*-algebras, this implication holds:

COROLLARY 5.2.4. *Let a short exact sequence of C^*-algebras*

$$0 \longrightarrow I \longrightarrow A \xrightarrow{\pi} F \longrightarrow 0$$

with finite-dimensional F be given. If I is subhomogeneous and semiprojective, then A is also semiprojective.

Proof. We verify condition (2) in Theorem 5.1.2 for A. By assumption, each $\operatorname{Prim}_k(I)$ is a one-dimensional ANR-space after compactification and the same holds for any space obtained from $\operatorname{Prim}_k(I)$ by adding finitely many points ([ST12, Theorem 6.1]). Hence the one-point compactifications of $\operatorname{Prim}_k(A)$ are 1-dimensional ANRs for all k. If we assume $F = \mathbb{M}_n$, then the set-valued retract maps R_k are unchanged for $k < n$. For $k = n$, regularity of R_k follows from regularity of the retract map for I and the fact that $\{[\pi]\}$ is closed in $\mathrm{Prim}_{\leq k}(A) = \mathrm{Prim}_{\leq k}(I) \cup \{[\pi]\}$. For $k > n$, we apply Lemma 4.1.6 to

$$\begin{tikzcd}
 & & 0 \arrow[d] & 0 \arrow[d] & \\
0 \arrow[r] & \mathcal{C}_0(\mathrm{Prim}_k(I), \mathbb{M}_k) \arrow[r] \arrow[d, equal] & I_{\leq k} \arrow[r] \arrow[d] & I_{\leq k-1} \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & \mathcal{C}_0(\mathrm{Prim}_k(A), \mathbb{M}_k) \arrow[r] & A_{\leq k} \arrow[r] \arrow[d, "\pi"] & A_{\leq k-1} \arrow[r] \arrow[d] & 0 \\
 & & F \arrow[r, equal] \arrow[d] & F \arrow[d] & \\
 & & 0 & 0 &
\end{tikzcd}$$

and see that $R_k \colon \mathrm{Prim}_{\leq k}(A) \to 2^{\mathrm{Prim}_{\leq k-1}(A)}$ is again lower semicontinuous and has pointwise finite image. □

5.2.2 Quantum permutation algebras

We are now going to demonstrate how the structure theorem 5.1.2 can be used to show that certain C^*-algebras fail to be semiprojective. We would like to thank T. Katsura for pointing out to us the quantum permutation algebras ([Wan98], [BC08]) as a testcase:

Definition 5.2.5 ([BC08]). *For $n \in \mathbb{N}$, the quantum permutation algebra $A_s(n)$ is the universal C^*-algebra generated by n^2 elements u_{ij}, $1 \leq i, j \leq n$, with relations*

$$u_{ij} = u_{ij}^* = u_{ij}^2 \quad \& \quad \textstyle\sum_j u_{ij} = \sum_i u_{ij} = 1.$$

It is not clear from the definition whether the C^*-algebras $A_s(n)$ are semiprojective or not. For $n \in \{1, 2, 3\}$ one easily finds $A_s(n) \cong \mathbb{C}^{n!}$ so that we have semiprojectivity in that cases. For higher n one might expect semiprojectivity of $A_s(n)$ because of the formal similarity to graph C^*-algebras. In fact, their definition only involves finitely many projections and orthogonality resp. order relations between them. Since graph C^*-algebras associated to finite graphs are easily seen to be semiprojective, one might think that we also have semiprojectivity for the quantum permutation algebras. This was even erroneously claimed to be true in [Bla04, example 2.8(vi)]. In this section we will show that the C^*-algebras $A_s(n)$ are in fact not semiprojective for all $n \geq 4$.
One can reduce the question for semiprojectivity of these algebras to the case $n = 4$. The following result of Banica and Collins shows that the algebra $A_s(4)$ is 4-subhomogeneous, so that our machinery applies. The idea is to get enough

information about the primitive spectrum of $A_s(4)$ to show that it contains closed subsets of dimension strictly greater than 1. This will then contradict part (2) of 5.1.2, so that $A_s(4)$ cannot be semiprojective.

We follow notations from [BC08] and denote the Pauli matrices by

$$c_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad c_2 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad c_3 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad c_4 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

Set $\xi_{ij}^x = c_i x c_j$ and regard $\mathbb{M}_2$ as a Hilbert space with respect to the scalar product $< a|b >= \operatorname{tr}(b^*a)$. Then for any $x \in SU(2)$ we find $\{\xi_{ij}^x\}_{j=1..4}$ and $\{\xi_{ij}^x\}_{i=1..4}$ to be a basis for $\mathbb{M}_2$. Under the identification $\mathbb{M}_4 \cong \mathcal{B}(\mathbb{M}_2)$, Banica and Collins studied the following representation of $A_s(4)$:

PROPOSITION 5.2.6 (Theorem 4.1 of [BC08]). *The *-homomorphism given by*

$$\begin{aligned} \pi\colon A_s(4) &\longrightarrow \mathcal{C}(SU(2), \mathbb{M}_4) \\ u_{ij} &\mapsto \big(x \mapsto \textit{rank one projection onto}\, \mathbb{C} \cdot \xi_{ij}^x\big) \end{aligned}$$

is faithful. It is called the Pauli representation of $A_s(4)$.

For the remainder of this section let S denote the following subset of $SU(2)$:

$$S := \left\{ \begin{pmatrix} \lambda & -\overline{\mu} \\ \mu & \overline{\lambda} \end{pmatrix} \in SU(2)\colon \min \left\{ \begin{array}{c} |\operatorname{Re}(\lambda\mu)|, |\operatorname{Im}(\lambda\mu)|, |\operatorname{Re}(\overline{\lambda}\mu)|, \\ |\operatorname{Im}(\overline{\lambda}\mu)|, ||\lambda| - |\mu|| \end{array} \right\} = 0 \right\}$$

We will now study the representations of $A_s(4)$ obtained by composing the Pauli representation with a point evaluation. As we will see, most points of $SU(2)$ lead to irreducible representations which are furthermore locally pairwise inequivalent.

LEMMA 5.2.7. *The representation $\pi_x = \operatorname{ev}_x \circ \pi\colon A_s(4) \to \mathbb{M}_4$ is irreducible for every $x \in SU(2)\backslash S$.*

Proof. Let $x = \begin{pmatrix} \lambda & -\overline{\mu} \\ \mu & \overline{\lambda} \end{pmatrix} \in SU(2)\backslash S$ be given, we show that the commutant of $\pi_x(A_s(4))$ equals the scalars. Therefore we will check the matrix entries of the elements $\pi_x(u_{ij})$ with respect to the orthonormal basis $\left\{ \frac{1}{\sqrt{2}}\xi_{11}^x, \frac{1}{\sqrt{2}}\xi_{12}^x, \frac{1}{\sqrt{2}}\xi_{13}^x, \frac{1}{\sqrt{2}}\xi_{14}^x \right\}$ of $M_2 \cong \mathbb{C}^4$. Since in this picture $\pi_x(u_{1i})$ equals the elementary matrix e_{ii}, every element in $(\pi_x(A_s(4)))'$ is diagonal. But we also find

$$\begin{aligned} (\pi_x(u_{23}))_{12} &= \tfrac{1}{2} < \pi_x(u_{23})\xi_{12}^x | \xi_{11}^x > \\ &= \tfrac{1}{4} < \xi_{12}^x | \xi_{23}^x >< \xi_{23}^x | \xi_{11}^x >= 4 \cdot \operatorname{Re}(\lambda\mu) Im(\lambda\mu) \neq 0, \\ (\pi_x(u_{22}))_{13} &= \tfrac{1}{4} < \xi_{13}^x | \xi_{22}^x >< \xi_{22}^x | \xi_{11}^x >= 2 \cdot \operatorname{Re}(\lambda\mu)(|\lambda|^2 - |\mu|^2) \neq 0, \\ (\pi_x(u_{22}))_{14} &= \tfrac{1}{4} < \xi_{14}^x | \xi_{22}^x >< \xi_{22}^x | \xi_{11}^x >= -2 \cdot \operatorname{Im}(\lambda\mu)(|\lambda|^2 - |\mu|^2) \neq 0. \end{aligned}$$

So the only elements of $\mathbb{M}_4$ commuting with all of $\pi_x(A_s(4))$ are the scalars. $\square$

Proposition 5.2.8. *Every $x \in SU(2)\backslash S$ admits a small neighborhood $V \subseteq SU(2)\backslash S$ such that for all distinct $y, y' \in V$ the representations π_y and $\pi_{y'}$ are not unitarily equivalent.*

Proof. Let $x = \begin{pmatrix} \lambda_0 & -\overline{\mu_0} \\ \mu_0 & \overline{\lambda_0} \end{pmatrix} \in SU(2)\backslash S$ be given, then

$$\epsilon := \min \left\{ \begin{array}{c} |\operatorname{Re}(\lambda_0\mu_0)|, |\operatorname{Im}(\lambda_0\mu_0)|, |\operatorname{Re}(\overline{\lambda_0}\mu_0)|, \\ |\operatorname{Im}(\overline{\lambda_0}\mu_0)|, \left||\lambda_0| - |\mu_0|\right|, |\lambda_0| \end{array} \right\} > 0.$$

Define a neighborhood $V \subseteq SU(2)\backslash S$ of x by

$$V = \left\{ \begin{pmatrix} \lambda & -\overline{\mu} \\ \mu & \overline{\lambda} \end{pmatrix} \in SU(2)\backslash S : |\lambda - \lambda_0| < \frac{\epsilon}{3}, |\mu - \mu_0| < \frac{\epsilon}{3} \right\}.$$

Now let $y, y' \in V$ with unitarily equivalent representations π_y and $\pi_{y'}$ be given. We compute the value

$$\begin{array}{ll} \|\pi_y(u_{11}u_{22})\| & = \frac{1}{4}\|(< - |\xi_{11}^y > \xi_{11}^y) \circ (< - |\xi_{22}^y > \xi_{22}^y)\| \\ & = \frac{1}{4}| < \xi_{22}^y|\xi_{11}^y > | \cdot \|(< - |\xi_{22}^y > \xi_{11}^y)\| \\ & = \frac{1}{4}| < \xi_{22}^y|\xi_{11}^y > | \cdot \|\xi_{22}^y\|\|\xi_{11}^y\| \\ & = \left||\lambda|^2 - |\mu|^2\right| \end{array}$$

which is invariant under unitary equivalence. So we find $\left||\lambda|^2 - |\mu|^2\right| = \left||\lambda'|^2 - |\mu'|^2\right|$. This implies

$$(|\lambda| = |\lambda'| \wedge |\mu| = |\mu'|) \quad \vee \quad (|\lambda| = |\mu'| \wedge |\mu| = |\lambda'|)$$

because of $|\lambda|^2 + |\mu|^2 = 1 = |\lambda'|^2 + |\mu'|^2$. By definition of V we have

$$\left||\lambda| - |\mu'|\right| \geq \left||\lambda_0| - |\mu_0|\right| - \left||\lambda| - |\lambda_0|\right| - \left||\mu'| - |\mu_0|\right| > \frac{\epsilon}{3} > 0,$$

so that we can exclude the second case. Analogously, computing the invariants $\|\pi_y(u_{13}u_{22})\|$ and $\|\pi_y(u_{14}u_{22})\|$ gives

$$|\operatorname{Re}(\lambda\mu)| = |\operatorname{Re}(\lambda'\mu')| \quad \text{and} \quad |\operatorname{Im}(\lambda\mu)| = |\operatorname{Im}(\lambda'\mu')|$$

and checking $\|\pi_y(u_{11}u_{42})\|$ and $\|\pi_y(u_{11}u_{32})\|$ shows

$$|\operatorname{Re}(\overline{\lambda}\mu)| = |\operatorname{Re}(\overline{\lambda'}\mu')| \quad \text{and} \quad |\operatorname{Im}(\overline{\lambda}\mu)| = |\operatorname{Im}(\overline{\lambda'}\mu')| \ .$$

The last four equalities imply $\lambda\mu = \lambda'\mu'$ and $\overline{\lambda}\mu = \overline{\lambda'}\mu'$ by the choice of V. Together with $|\lambda| = |\lambda'|$ and $|\mu| = |\mu'|$ we find $(\lambda, \mu) = (\lambda', \mu')$ or $(\lambda, \mu) = (-\lambda', -\mu')$. In the second case we get $|\lambda - \lambda'| = 2|\lambda| \geq 2|\lambda_0| - 2|\lambda - \lambda_0| \geq \frac{4\epsilon}{3}$ contradicting $|\lambda - \lambda'| \leq |\lambda - \lambda_0| + |\lambda' - \lambda_0| < \frac{2\epsilon}{3}$ by the choice of V. It follows that $y = y'$. $\square$

By now we have obtained enough information about $\operatorname{Prim}(A_s(4))$ to show that it does not satisfy condition (2) of Theorem 5.1.2. Hence we find:

Theorem 5.2.9. *The C^*-algebra $A_s(4)$ is not semiprojective.*

Proof. Choose a point $x_0 \in SU(2)\backslash S$ and a neighborhood V of x_0 as in Proposition 5.2.8. Since $SU(2)$ is a real 3-manifold, there is a neighborhood of x_0 contained in V which is homeomorphic to $\mathbb{D}^3 = \{x \in \mathbb{R} \colon \|x\| \leq 1\}$. The restriction of the Pauli representation π to this neighborhood gives a *-homomorphism $\varphi\colon A_s(4) \to \mathcal{C}(\mathbb{D}^3, \mathbb{M}_4)$ with the property that $\mathrm{ev}_x \circ \varphi$ and $\mathrm{ev}_y \circ \varphi$ are irreducible but not unitarily equivalent for all distinct $x, y \in \mathbb{D}^3$. The pointwise surjectivity of φ given by Lemma 5.2.7 and a Stone-Weierstraß argument ([Kap51, Theorem 3.1]) show that φ is in fact surjective. This implies that $\mathrm{Prim}_4(A_s(4))$ contains a closed 3-dimensional subset and hence $\dim(\mathrm{Prim}_4(A_s(4))) \geq 3$. As a consequence, $A_s(4)$ cannot be semiprojective because it is subhomogeneous by Proposition 5.2.6 but fails to satisfy condition (2) of Theorem 5.1.2. $\square$

It is not hard to show that semiprojectivity of $A_s(n)$ for some $n > 4$ would force $A_s(4)$ to be semiprojective. Since we have just shown that this is not the case, we obtain:

Corollary 5.2.10. *The C^*-algebras $A_s(n)$ are not semiprojective for $n \geq 4$.*

Proof. For $n \geq 4$ there is a canonical surjection $\varrho_n\colon A_s(n) \to A_s(4)$ given by

$$u_{ij}^{(n)} \mapsto \begin{cases} u_{ij}^{(4)} & \text{if } 1 \leq i,j \leq 4 \\ 1 & \text{if } i = j > 4 \\ 0 & \text{otherwise} \end{cases} .$$

The kernel of ϱ_n is generated by the finite set of projections $\left\{u_{ij}^{(n)} \colon \varrho_n\left(u_{ij}^{(n)}\right) = 0\right\}$. It follows from [Sør12, Proposition 3], which extends the idea of [Neu00, Proposition 5.19], that semiprojectivity of $A_s(n)$ would imply semiprojectivity of $\varrho_n(A_s(n)) = A_s(4)$. Since this is not the case by Theorem 5.2.9, $A_s(n)$ cannot be semiprojective for all $n \geq 4$. $\square$

References

[AF90] J.-P. Aubin and H. Frankowska. *Set-valued analysis*, Birkhäuser, Boston, 1990.

[BC08] T. Banica and B. Collins. Integration over the Pauli quantum group, *J. Geom. Phys.*, 58(8):942-961, 2008.

[BH74] J.J. Bastian and K.J. Harrison. Subnormal weighted shifts and asymptotic properties of normal operators, *Proc. Amer. Math. Soc.*, 42(2):475-479, 1974.

[Bin49] R.H. Bing. Partitioning a set, *Bull. Amer. Math. Soc.*, 55:1101-1110, 1949.

[Bla85] B. Blackadar. Shape theory for C^*-algebras, *Math. Scan.*, 56(2):249-275, 1985.

[Bla04] B. Blackadar. Semiprojectivity in simple C^*-algebras, in *Operator algebras and applications*, volume 38 of *Adv. Stud. Pure Math.*, pages 1-17. Math. Soc. Japan, Tokyo, 2004.

[Bla06] B. Blackadar. *Operator algebras*, volume 122 of *Encyclopedia of Mathematical Sciences*, Springer-Verlag, Berlin, 2006.

[Bla12] B. Blackadar. The homotopy lifting theorem for semiprojective C^*-algebras, arXiv preprint math.OA/1207.1909v3, 2012.

[Bor67] K. Borsuk. *Theory of retracts*, Monografie Matematyczne, 44, Warszawa: PWN - Polish Scientific Publishers, 1967.

[Bro00] N.P. Brown. On quasidiagonal C^*-algebras, arXiv preprint math.OA/0008181, 2000.

[CD10] A. Chigogidze and A.N. Dranishnikov. Which compacta are noncommutative ARs?, Topology Appl. 157(4):774-778, 2010.

[Dix77] J. Dixmier. *C^*-algebras*, North-Holland, Amsterdam and New York, 1977.

[ELP98] S. Eilers, T. A. Loring and G. K. Pedersen. Stability of anticommutation relations: An application of noncommutative CW complexes, *J. reine angew. Math.*, 499:101-143, 1998.

[ELP99] S. Eilers, T. A. Loring and G. K. Pedersen. Morphisms of extensions of C^*-algebras: Pushing forward the Busby invariant, *Advances in Mathematics*, 147:74-109, 1999.

[End14] D. Enders. Subalgebras of finite codimension in semiprojective C^*-algebras, arXiv preprint math.OA/1405.2750, 2014.

[EK] S. Eilers and T. Katsura. Semiprojectivity and properly infinite projections in graph C^*-algebras, in preparation.

[Fel61] J.M.G. Fell. The structure of algebras of operator fields, *Acta Math.* 106:233-280, 1961.

[FR01] P. Friis and M. Rørdam. Approximation with normal operators with finite spectrum and an elementary proof of a Brown-Douglas-Fillmore theorem, *Pacific J. Math.*, 199(2):347-366, 2001.

[GP84] K. Grove and G.K. Pedersen. Sub-stonean spaces and corona sets, *Journal of Functional Analysis*, 56:124-143, 1984.

[HW48] W. Hurewicz and H. Wallman, *Dimension theory*, Princeton University Press, Princeton, 1948.

[Kap51] I. Kaplansky. The structure of certain operator algebras, *Trans. Amer. Math. Soc.*, 70:219-255, 1951.

[Lin97] H. Lin. Almost commuting selfadjoint matrices and applications, *Operator algebras and their applications* (Waterloo, ON, 1994/1995), 193-233, Fields Inst. Commun. 13, American Mathematical Society, Providence, RI, 1997.

[Lor97] T.A. Loring. *Lifting solutions to perturbing problems in C^*-algebras*, volume 8 of *Fields Institute Monographs*, American Mathematical Society, Providence, RI, 1997.

[LP98] T.A. Loring and G.K. Pedersen. Projectivity, transitivity and AF-telescopes, *Trans. Amer. Math. Soc.*, 350(11):4313-4339, 1998.

[Mei05] M. Meilstrup. Classifying homotopy types of one-dimensional Peano continua, Thesis for Master of Science at the Department of Mathematics at Brigham Young University, 2005.

[Moi49] E. Moise, Grille decomposition and convexification theorems for compact metric locally connected continua, *Bull. Amer. Math. Soc*, 55:1111-1121.

[Nad92] S.B.jun. Nadler. *Continuum theory: An introduction*, volume 158 of *Pure and Applied Mathematics*, Marcel Dekker, New York, 1992.

[Nag70] K. Nagami. *Dimension Theory*, Pure Appl. Math., No. 37, Academic Press, London, New York, 1970.

[Neu00] B. Neubüser. Semiprojektivität und Realisierungen von rein unendlichen C^*-Algebren, Ph.D. thesis, Universität Münster, 2000.

[Phi07] N.C. Phillips, Recursive subhomogeneous algebras, *Trans. Amer. Math. Soc.*, 359:4595-4623, 2007.

[RLL00] M. Rørdam, F. Larsen and N. Lautsen. *An introduction to K-theory for C^*-algebras*, volume 49 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 2000.

[Rør02] M. Rørdam. Classification of nuclear, simple C^*-algebras, in *Classification of nuclear C^*-algebras. Entropy in operator algebras*, volume 126 of *Encyclopaedia Math. Sci.*, pages 1-145, Springer, Berlin, 2002.

[Sør12] A.P.W. Sørensen. On a counterexample to a conjecture by Blackadar, Operator Algebra and Dynamics, Springer Proceedings in Mathematics & Statistics, Vol. 58, 2013.

[RW98] I. Raeburn and D.P. Williams. *Morita equivalence and continuous-trace C^*-algebras*, vol. 60 of *Mathematical Surveys and Monographs*, AMS, Providence, RI, 1998.

[ST12] A.P.W. Sørensen and H. Thiel, A characterization of semiprojectivity for commutative C^*-algebras, *Proc. Lond. Math. Soc.*, 105(5):1021-1046, 2012.

[Wan98] S. Wang. Quantum symmetry groups of finite spaces, *Comm. Math. Phys.*, 195:195-211, 1998.

[War60] L.E. Ward. On local trees, *Proc. Amer. Math. Soc.*, 11:940-944, 1960.

[WO93] N.E. Wegge-Olsen. *K-Theory and C^*-algebras: a friendly approach*, Oxford University Press, 1993.

Dominic Enders
University of Copenhagen
Department of
Mathematical Sciences
Universitetsparken 5
2100 Copenhagen, Denmark
d.enders@math.ku.dk